Titre: Title:	Sécurité dans les réseaux mobiles de nouvelle génération
Auteur: Author:	Angelo Rossi
Date:	2011
Type:	Mémoire ou thèse / Dissertation or Thesis
Référence: Citation:	Rossi, A. (2011). Sécurité dans les réseaux mobiles de nouvelle génération [Thèse de doctorat, École Polytechnique de Montréal]. PolyPublie. https://publications.polymtl.ca/614/

URL de PolyPublie: PolyPublie URL:	https://publications.polymtl.ca/614/
Directeurs de recherche: Advisors:	Samuel Pierre
Programme: Program:	Génie informatique

UNIVERSITÉ DE MONTRÉAL

SÉCURITÉ DANS LES RÉSEAUX MOBILES DE NOUVELLE GÉNÉRATION

ANGELO ROSSI
DÉPARTEMENT DE GÉNIE INFORMATIQUE ET GÉNIE LOGICIEL
ÉCOLE POLYTECHNIQUE DE MONTRÉAL

THÈSE PRÉSENTÉE EN VUE DE L'OBTENTION
DU DIPLÔME DE PHILOSOPHIÆ DOCTOR
(GÉNIE INFORMATIQUE)
MAI 2011

UNIVERSITÉ DE MONTRÉAL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

Cette thèse intitulée :

SÉCURITÉ DANS LES RÉSEAUX MOBILES DE NOUVELLE GÉNÉRATION

présentée par : ROSSI, Angelo
en vue de l'obtention du diplôme de : Philosophiæ Doctor
a été dûment acceptée par le jury d'examen constitué de :

Mme. BELLAICHE, Martine, Ph.D., présidente.
M. PIERRE, Samuel, Ph.D., membre et directeur de recherche.
M. QUINTERO, Alejandro, Doct., membre.
M. ROBERT, Jean-Marc, Ph.D., membre externe.

À Catherine, mon amour inconditionnel ;
à mes parents et ma soeur ;
à mes collègues du laboratoire LARIM.

REMERCIEMENTS

Avoir l'appui d'un directeur de recherche dynamique qui se soucie des besoins de ses étudiants autant du point de vue technique, moral que financier est certainement un atout pour tout étudiant aux cycles supérieures. Pour cette raison, mon premier remerciement va à mon directeur de recherche Samuel Pierre qui, par ses nombreux conseils, m'a permis d'approfondir ma passion innée : la sécurité des réseaux informatiques.

J'aimerais également remercier les membres du département de recherche et développement de Ericsson Research Canada en particulier Denis Monette, Laurent Marchand, Suresh Krishnan, Makan Pourzandi et Frédéric Rossi. Leur motivation et leur vision avant-gardiste dans le domaine des télécommunications m'ont su être utiles pour concevoir des solutions sécuritaires innovatrices qui se sont finalisées par le dépôt de 5 brevets.

Merci aussi aux membres du laboratoire de recherche en informatique mobile (LARIM) que j'ai eu la chance de côtoyer tout au long de mes études aux cycles supérieures. Leur enthousiasme et leur plaisir d'aider ont fait du laboratoire un endroit des plus agréables à y vivre et travailler. Une mention spéciale doit être portée à mes collègues Georges Abou-Khalil et Stéphane Ouellette qui sont devenus au fil des années des amis proches avec qui j'espère entretenir une relation à long terme.

Un gros merci également aux fonds québécois de la recherche sur la nature et les technologies (FQRNT) et au conseil de recherches en sciences naturelles et en génie du Canada (CRSNG) qui m'ont donnée un appui financier important via la bourse BMP innovation afin que je puisse me concentrer entièrement sur mes études.

Je remercie tous les professeurs du département du génie informatique, notamment Guillaume-Alexandre Bilodeau, Michel Gagnon et Michel Dagenais pour m'avoir donné l'opportunité d'enseigner. Il n'y a rien de plus valorisant pour moi que de partager mon savoir avec des étudiants motivés d'apprendre.

Mes remerciements s'étendent aussi aux membres du jury également pour la révision et l'évaluation de cette thèse.

À un niveau plus personnel, le support de ma famille est incontestablement la raison principale qui m'a permis de me surpasser et de terminer mes études doctorales. Sans vous je ne saurais trouver le courage de persévérer dans les moments plus difficiles. Je ne pourrai terminer mes remerciements sans souligner la patience, le support moral et l'amour inconditionnel que me donne ma douce moitié, Catherine Leroux. Si ce n'était de ta compréhension, ma chérie, cet ouvrage n'aurait jamais atteint la qualité que je voulais y attribuer. Cette thèse ne représente qu'une infime réalisation parmi celles que nous accomplieront ensemble.

RÉSUMÉ

Les réseaux de nouvelle génération visent à converger les réseaux fixes et mobiles hétérogènes afin d'offrir tous les services à travers un réseau coeur tout IP. Faisant parti du réseau d'accès mobile, un des principaux objectifs du réseau 4G est de permettre une relève ininterrompue entre les réseaux cellulaires et WIFI pour ainsi favoriser l'apprivoisement de services vidéo mobiles exigeant des critères de qualité de service très stricts à moindres coûts. Cependant, l'uniformisation du trafic au niveau de la couche réseau favorise sa centralisation à travers un réseau coeur IP partagé par tous les opérateurs, la rendant ainsi comme une cible vulnérable de choix pour les pirates informatiques. La conception de solutions sécuritaires dans un environnement où les entités ne se connaissent pas à priori s'annonce comme une tâche très ardue.

La thèse se penche sur quatre problématiques importantes dans les réseaux de nouvelle génération dont chacune est traitée dans un article distinct. Les deux premiers articles touchent à la sécurité dans un contexte décentralisé, à savoir les réseaux mobiles ad hoc (MANETs), alors que les deux derniers proposent des mécanismes innovateurs pour sécuriser des solutions visant à réduire la consommation de bande passante et d'énergie, en conformité avec le virage vert informatique promu par les opérateurs réseautiques. Plus précisément, le troisième article traite de la sécurisation des flots multicast dans un environnement à haut taux de perte de paquet et le dernier propose une solution d'optimisation de route sécuritaire pour mobile IPv6 (MIPv6) utilisant une version améliorée de l'algorithme de génération d'adresses cryptographiques (CGA) et les extensions de sécurité du système de nom de domaine (DNSSEC).

Les systèmes de détection d'intrusion (IDS) pour les MANETs basés sur la réputation des nœuds classifient les participants du réseau selon leur degré de confiance. Cependant, ils partagent tous une vulnérabilité commune : l'impossibilité de détecter et de réagir aux attaques complices. Le premier article propose un IDS qui intègre efficacement le risque de collusion entre deux ou plusieurs nœuds malveillants dans le calcul de la fiabilité d'un chemin. L'algorithme proposé ne se limite pas qu'au nombre et à la réputation des nœuds intermédiaires formant un chemin, mais intègre également d'autres informations pertinentes sur les voisins des nœuds intermédiaires d'un chemin pouvant superviser le message original et celui retransmis. Le IDS proposé détecte efficacement les nœuds malicieux et complices dans le but de les isoler rapidement du réseau. Les simulations lancées dans divers environnements MANETs contenant une proportion variable d'attaquants complices montrent bien l'efficacité du IDS proposée en offrant un gain en débit considérable comparativement aux solutions existantes.

À l'instar de prévenir les comportements égoïstes des nœuds par la menace d'être privés de certaines fonctions, voire même isolés du réseau, due à une baisse de réputation, le second article opte pour un incitatif non-punitif en la monnaie virtuelle plus communément appelée nuglets. Plus précisément, l'article présente un cadre de travail issu de la théorie des jeux basé sur la compétition de Bertrand pour inciter les nœuds intermédiaires à retransmettre les messages selon les requis de QoS demandés par la source. Pour qu'un nœud source envoie ou accède à un flot sensible à la QoS comme par exemple les applications en temps réel, il débute par envoyer un contrat qui spécifie les critères de QoS, sa durée et son prix de réserve. Sur réception du contrat, les nœuds intermédiaires formant une route entre la source et la destination partagent les informations sur eux-mêmes et celles recueillies sur les nœuds voisins, anciens et courants pour estimer la probabilité de bris de contrat ainsi que le nombre de compétiteurs actifs. Ces deux paramètres sont cruciaux dans le processus de fixation des prix. Une fois les réponses de route recueillies, la source choisit la route la moins chère. Le cadre de travail multijoueur proposé, basé sur la compétition de Bertrand avec des firmes asymétriques et ayant accès à de l'information imparfaite, possède un équilibre de Nash en stratégies mixtes dans lequel le profit des firmes est positif et baisse non seulement avec le nombre de compétiteurs, mais aussi avec l'impression d'une précision accrue que les compétiteurs ont sur le coût de production du joueur. Les résultats montrent que l'incertitude sur les coûts augmente le taux de la marge brute et la fluctuation des prix tout en diminuant les chances d'honorer le contrat.

Dans un autre ordre d'idée, l'intérêt sans cesse grandissant des opérateurs à converger les réseaux fixes et mobiles dans le but d'offrir une relève sans interruption favorise l'utilisation des applications vidéo mobiles qui surchargeront rapidement leurs réseaux. Dans un contexte du virage vert qui prend de plus en plus d'ampleur dans le domaine des télécommunications, la transmission des flots en multidiffusion (multicast) devient essentiel dans le but de réduire la consommation de bande passante et la congestion du réseau en rejoignant simultanément plusieurs destinataires. La sécurisation des flots en multidiffusion a été largement étudiée dans la littérature antérieure, cependant aucune des solutions proposées ne tient compte des contraintes imposées par les liaisons sans fil et la mobilité des nœuds, en particulier le haut taux de perte de paquets. La nécessité d'un mécanisme de distribution de clés régénératrices efficace et pouvant supporter un grand bassin d'abonnés pour les réseaux mobiles n'aura jamais été aussi urgent avec l'arrivée de la convergence fixe-mobile dans les réseaux 4G. Le troisième article présente deux algorithmes de clés régénératrices basés sur les chaînes de hachage bidirectionnelles pour le protocole de distribution de clés logical key hierarchy (LKH). Ainsi, un membre ayant perdu jusqu'à un certain nombre de clés de déchiffrement consécutives pourrait lui-même les régénérer sans faire la requête de retransmission au serveur de clés.

Les simulations effectuées montrent que les algorithmes proposés offrent des améliorations considérables dans un environnement de réseau mobile à taux de perte de paquet, notamment dans le percentage de messages déchiffrés.

Le souci d'efficacité énergétique est également présent pour les opérateurs de réseaux cellulaires. D'ailleurs, près de la moitié des abonnements sur Internet proviennent présentement d'unités mobiles et il est attendu que ce groupe d'utilisateurs deviennent le plus grand bassin d'usagers sur Internet dans la prochaine décennie. Pour supporter cette croissance rapide du nombre d'utilisateurs mobiles, le choix le plus naturel pour les opérateurs serait de remplacer mobile IPv4 par MIPv6. Or, la fonction d'optimisation de route (RO), qui remplace le routage triangulaire inefficace de MIP en permettant au nœud mobile (MN) une communication bidirectionnelle avec le nœud correspondant (CN) sans faire passer les messages à travers l'agent du réseau mère (HA), est déficiente au niveau de la sécurité. L'absence d'informations pré-partagées entre le MN et le CN rend la sécurisation du RO un défi de taille. MIPv6 adopte la routabilité de retour (RR) qui est davantage un mécanisme qui vérifie l'accessibilité du MN sur son adresse du réseau mère (HoA) et du réseau visité (CoA) plutôt qu'une fonction de sécurité. D'autres travaux se sont attaqués aux nombreuses failles de sécurité du RR, mais soit leur conception est fautive, soit leurs suppositions sont irréalistes. Le quatrième article présente une version améliorée de l'algorithme de génération cryptographique d'adresse (ECGA) pour MIPv6 qui intègre une chaîne de hachage arrière et offre de lier plusieurs adresses CGA ensemble. ECGA élimine les attaques de compromis temps-mémoire tout en étant efficace. Ce mécanisme de génération d'adresse fait parti du protocole Secure MIPv6 (SMIPv6) proposé avec un RO sécuritaire et efficace grâce à DNSSEC pour valider les CGAs qui proviennent d'un domaine de confiance et qui permet une authentification forte plutôt que l'invariance de source. Le vérificateur de protocoles cryptographiques dans le modèle formel *AVISPA* a été utilisé pour montrer qu'aucune faille de sécurité n'est présente tout en limitant au maximum les messages échangés dans le réseau d'accès.

ABSTRACT

Next generation networks aim at offering all available services through an IP-core network by converging fixed-mobile heterogeneous networks. As part of the mobile access network, one of the main objectives of the 4G network is to provide seamless roaming with wireless local area networks and accommodating quality of service (QoS) specifications for digital video broadcasting systems. Such innovation aims expanding video-based digital services while reducing costs by normalizing the network layer through an all-IP architecture such as Internet. However, centralizing all traffic makes the shared core network a vulnerable target for attackers. Design security solutions in such an environment where entities a priori do not know each other represent a daunting task.

This thesis tackles four important security issues in next generation networks each in distinct papers. The first two deal with security in decentralized mobile ad hoc networks (MANETs) while the last two focus on securing solutions aiming at reducing bandwidth and energy consumption, in line with the green shift promoted by network operators. More precisely, the third paper is about protecting multicast flows in a packet-loss environment and the last one proposes a secure route optimization function in mobile IPv6 (MIPv6) using an enhanced version of cryptographically generated address (CGA) and domain name service security extensions (DNSSEC).

Most intrusion detection systems (IDS) for MANETs are based on reputation system which classifies nodes according to their degree of trust. However, existing IDS all share the same major weakness: the failure to detect and react on colluding attacks. The first paper proposes an IDS that integrates the colluding risk factor into the computation of the path reliability which considers the number and the reputation of nodes that can compare both the source message and the retransmitted one. Also, the extended architecture effectively detects malicious and colluding nodes in order to isolate them and protect the network. The simulations launched in various MANETs containing various proportions of malicious and colluding nodes show that the proposed solution offers a considerable throughput gain compared to current solutions. By effectively selecting the most reliable route and by promptly detecting colluding attacks, the number of lost messages is decreased, and therefore, offering more efficient transmissions.

Instead of thwarting selfishness in MANETs by threatening nodes to limit their network functions, the second paper opts for a non-punishment incentive by compensating nodes for their service through the use of virtual money, more commonly known as nuglets. The last paper presents a game-theoretic framework based on Bertrand competition to incite relaying

nodes in forwarding messages according to QoS requirements. For a source to send or access QoS-sensitive flows, such as real-time applications, it starts by sending a contract specifying the QoS requirements, its duration and a reservation price. Upon receiving a contract submission, intermediary nodes forming a route between the source and the destination share their current and past collected information on themselves and on surrounding nodes to estimate the probability of breaching the contract and the number of active competitors. Both parameters are crucial in setting a price. Once the source gets the responses from various routes, it selects the most cheapest one. This multiplayer winner-takes-all framework based on Bertrand competition with firms having asymmetric costs and access imperfect information has a mixed-strategy equilibrium in which industry profits are positive and decline not only with the number of firms having an estimated cost below the reservation price but also with the perception of a greater accuracy on a player's cost that competitors have. In fact, results show that cost uncertainty increases firms' gross margin rate and the prices fluctuation while making the contract honoring much riskier.

On another topic, with the growing interest in converging fixed and mobile networks, mobile applications will require more and more resources from both the network and the mobile device. In a social-motivated context of shifting into green technologies, using multicast transmissions is essential because it lowers bandwidth consumption by simultaneously reaching a group of multiple recipients. Securing multicast flows has been extensively studied in the past, but none of the existing solutions were meant to handle the constraints imposed by mobile scenarios, in particular the high packet-loss rate. The need for a low overhead self-healing rekeying mechanism that is scalable, reliable and suitable for mobile environments has never been more urgent than with the arrival of fixed-mobile convergence in 4G networks. The second paper presents two self-healing recovery schemes based on the dual directional hash chains for the logical key hierarchy rekeying protocol. This enables a member that has missed up to m consecutive key updates to recover the missing decryption keys without asking the group controller key server for retransmission. Conducted simulations show considerable improvements in the ratio of decrypted messages and in the rekey message overhead in high packet loss environments.

The concern of energy efficiency is also present for mobile access network operators. In fact, nearly half of all Internet subscribers come from mobile units at the moment and it is expected to be the largest pool of Internet users by the next decade. The most obvious choice for mobile operators to support more users would be to replace Mobile IP for IPv4 with MIPv6. However, the Route Optimization (RO) function, which replaces the inefficient triangle routing by allowing a bidirectional communication between a mobile node (MN) and the corresponding node (CN) without passing through its home agent (HA), is not secure and

has a high overhead. The lack of pre-shared information between the MN and the CN makes security in RO a difficult challenge. MIPv6 adopts the return routability (RR) mechanism which is more to verify the MN reachability in both its home address (HoA) and care-of address (CoA) than a security feature. Other works attempted to solve the multiple security issues in RR but either their design are flawed, or rely on unrealistic assumptions. The third paper presents an enhanced cryptographically generated address (ECGA) for MIPv6 that integrates a built-in backward key chain and offers support to bind multiple logically-linked CGAs together. ECGA tackles the time-memory tradeoff attacks while being very efficient. It is part of the proposed secure MIPv6 (SMIPv6) with secure and efficient RO which uses DNSSEC to validate CGAs from trusted domains and provide strong authentication rather than sender invariance. The AVISPA on-the-fly model checker (OFMC) tool has been used to show that the proposed solution has no security flaws while still being lightweight in signalling messages in the radio network.

TABLE DES MATIÈRES

LISTE DES TABLEAUX

LISTE DES FIGURES

LISTE DES SIGLES ET ABRÉVIATIONS

4G	Fourth Generation
AH	Authentication Header
AVISPA	Automated Validation of Internet Security Protocols and Applications
BT	Bidirectional Tunnelling
BU	Binding Update
CAM	Child-proof Authentication for MIPv6
(H)CBU	(Hierarchical) Certificate-based Binding Update
(E)CGA	(Enhanced) Cryptographically Generated IPv6 Address
CN	Correspondant Node
CONFIDANT	Automated Validation of Internet Security Protocols and Applications
CORE	Collaborative Reputation mechanism
CoT(I)	Care-of Token (Initialization)
DDHC	Dual Directional Hash Chain
DH	Diffie-Hellman
DNS(SEC)	Domain Name System Security Extensions
(D)DoS	(Distributed) Denial of Service
DSR	Dynamic Source Routing
ELK	Efficient Large-group Key
ESP	Encapsulating Security Payload
(W)FEC	(Weighted) Forward Erasure Correction
FMC	Fixed Mobile Convergence
FQDN	Fully Qualified Domain Name
GCKS	Group Controller Key Server
GMR	Gross Margin Rate
(V)HA	(Visited) Home Agent
HLPSL	High-Level Protocol Specification Language
HLSP	Home Link Subnet Prefix
HoT(I)	Home Token (Initialization)
IDS	Intrusion Detection System
IETF	Internet Engineering Task Force
KEK	Key Encryption Key
LKH	Logical Key Hierarchy
MANETs	Mobile Ad hoc Networks

MDS	Maximum Distance Seperable
(F)(H)(O)(P)(S)MIP	(Fast) (Hierarchical) (Optimizing) (Proxy) (Secure) Mobile IP
MITM	Man-In-The-Middle
MN	Mobile Node
OFT	One-way Function Tree
PRF	Pseudo Random Function
QoS	Quality of Service
RFC	Request For Comments
RO	Routing Optimization
RR	Return Routability
RREP	Route Reply
RREQ	Route Request
RSE	Reed-Solomon Erasure
SA	Security Association
SDR	Subset Difference Revocation
WKA	Weighted Key Assignment

CHAPITRE 1

INTRODUCTION

La convergence des réseaux hétérogènes fixes et mobiles visent une uniformisation globale afin de réduire les coûts ainsi que de rendre plus accessibles les services interactifs et vidéo tout en assurant leur continuité dans les relèves verticales à travers les différents réseaux et médias. Puisque cette évolution technologique mondiale ne peut être assurée que par la centralisation des flots de données à travers un réseau coeur partagé par tous les opérateurs, à savoir Internet, les pirates informatiques ont alors une cible d'attaque de choix sur laquelle porter leurs attaques. Le défi est d'autant plus grand avec le mouvement du virage vert qu'entreprend l'industrie des télécommunications et les ressources très limitées des unités mobiles en contraste avec les applications vidéo de plus en plus gourmandes qui poussent la recherche à concevoir des solutions sécuritaires efficaces.

Cette thèse porte sur trois aspects importants des réseaux de prochaine génération. D'abord, les réseaux mobiles ad hoc (MANETs), considérés comme une des technologies les plus prometteuses il y a une décennie, a vu son développement ralentir considérablement suite à la complexité de mettre en place un protocole de routage sécuritaire qui tient compte de l'indépendance des nœuds et du système décentralisé et distribué. En effet, contrairement aux réseaux classiques où la fonction de routage est assurée par des entités de confiance, cette fonction critique est prise en charge par les nœuds qui ne servent que leur intérêt individuel plutôt que collectif. Les services vidéo exigent une bande passante importante engendrant des coûts énormes aux opérateurs de réseaux. Dans ce contexte, les flots multidiffusés (multicast) vont devenir un enjeu important de compétitivité. Ainsi, le second volet de la thèse traite de la sécurisation du trafic d'une source vers plusieurs destinataires dans un environnement où la qualité du signal radio est instable. Toujours dans l'optique de l'efficacité énergétique et de la réduction de la bande passante, le mécanisme d'optimisation de route (RO) dans mobile IPv6 (MIPv6) présente des failles importantes compromettant l'adoption du protocole comme remplacement à MIPv4 dû à l'accroissement exponentiel du nombre d'usagers mobiles à s'abonner dans les années à venir.

Ce chapitre d'introduction présente d'abord quelques concepts de base qui serviront d'ancrage à la compréhension des sujets abordés. Suivront ensuite, les éléments de la problématique, les objectifs de recherche qui en découlent et la méthodologie adoptée. Pour terminer, les principales contributions de cette thèse et leur originalité seront discutées alors qu'une esquisse des autres chapitres clôturera le chapitre.

1.1 Définitions et concepts de base

Cette section est réservée à l'introduction de quelques concepts utiles dans la compréhension des solutions existantes et de celles proposées. La première partie se penche sur la définition des réseaux informatiques mobiles et du mode de fonctionnement des protocoles de base. Les réseaux ad hoc, la multidiffusion du trafic et le protocole MIPv6 avec ses modes du tunnel bidirectionnel (BT) et de la route optimisée seront donc introduits. La seconde moitié présente les attaques les plus populaires portées dans les réseaux ainsi que les objectifs de sécurité de base avec quelques primitives de sécurité pour les réaliser.

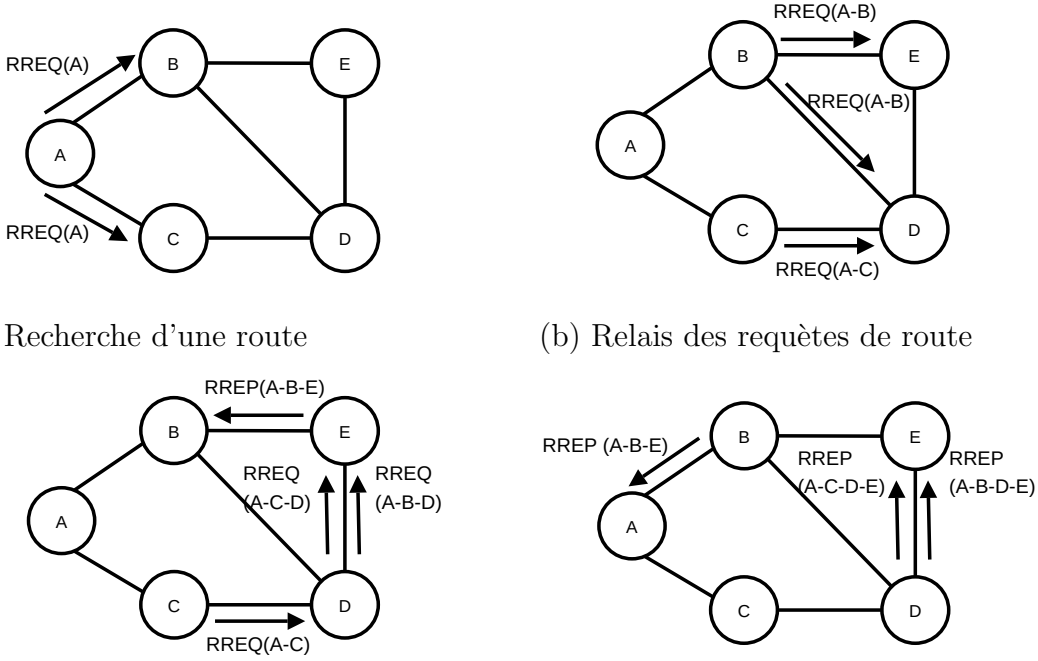
1.1.1 Réseaux informatiques mobiles et protocoles

Réseaux mobiles ad hoc (MANETs)

Les réseaux mobiles ad hoc est un regroupement d'unités mobiles dotés d'au moins une interface réseau sans fil permettant de communiquer avec d'autres nœuds sans infrastructure central. Les nœuds du réseau doivent donc assurer à la fois le rôle de routeur, permettant d'acheminer les messages de manière autonome, et de station hôte, pour recevoir et initier l'envoi de messages. Le réseau ad hoc se distingue des réseaux classiques par leur topologie très dynamique et aléatoire et par les ressources limitées poussant les nœuds à se comporter de manière égoïste. La Figure 1.1 illuste la procédure typique de recherche de route par lequel une source voulant rejoindre une destination débute par envoyer une requête de route que les nœuds intermédiaires retransmettent jusqu'à atteindre la destination. Cette dernière répond alors à l'aide d'une réponse de route qui revient jusqu'à la source et qui lui indique le chemin à utiliser pour envoyer ses données.

Multidiffusion

La multidiffusion permet une communication un-à-plusieurs par lequel une source rejoint simultanément un groupe de destinataires à l'aide d'un seul message que les routeurs du réseau, supportant la transmission diffusée sélective, peuvent copier et envoyer dans un ou plusieurs de leurs ports (voir Figure 1.2). Dans le any-source-multicast, plusieurs sources peuvent émettre leurs messages dans le même groupe et un nœud y adhère en précisant simplement l'adresse multicast qu'il désire s'abonner. La simplicité de cette méthode cache cependant la difficile tâche de la découverte des sources que doivent effectuer les routeurs ainsi que le manque inhérent de contrôle d'admission permettant à tous les nœuds d'envoyer des données. Au contraire, le single-source multicast force les abonnées à préciser non seulement l'adresse multicast, mais également l'identité de la source des flots qu'ils désirent recevoir.



(a) Recherche d'une route

(b) Relais des requètes de route

(c) Émission des réponses de route

(d) Route trouvée

Figure 1.1 Exemple de fonctionnement du protocole de routage DSR.

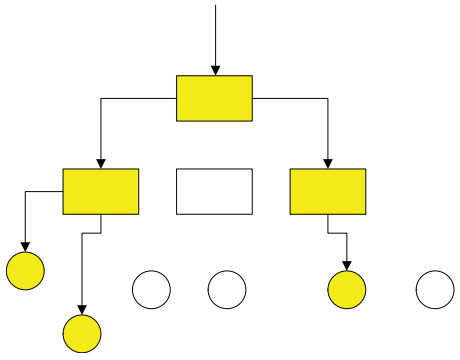


Figure 1.2 Exemple d'une multidiffusion

Mobile IPv6 (MIPv6)

Mobile IP est une extension de mobilité sur la couche IP afin pour préciser des champs importants pour permettre, entre autres, aux nœuds mobiles (MN) de lier son adresse du réseau visité (CoA) avec celui de son réseau mère (HoA). La forte croissance du nombre d'abonnés mobiles pousse les opérateurs de réseaux mobiles à envisager une alternative à MIPv4 qui offre un nombre d'adresses IP trop limité. MIPv6 devient donc un candidat naturel pour les opérateurs qui désirent garder les opérations de signalisation toujours initiées par le MN. Tout comme MIPv4, le MN possède 2 adresses IP dont celle du réseau mère est permanente alors que celle du réseau visité est temporaire, conservant ainsi le tunnel bidirectionnel (BT) dans lequel tout le trafic passe par l'agent du réseau mère (HA)(Figure 1.3). Cependant

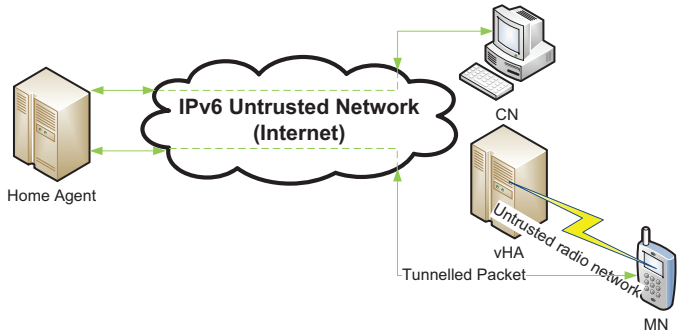


Figure 1.3 Tunnel bidirectionnel dans MIPv6

MIPv6 n'offre pas le routage triangulaire par lequel uniquement les données échangées en amont entre le MN et le nœud correspondant (CN) doivent obligatoirement traverser le réseau mère. Au contraire, le protocole permet une communication directe et optimale entre le MN et le CN (Figure 1.4), réduisant considérablement la surcharge vers le réseau mère ainsi que les délais. La réticence des opérateurs à adopter MIPv6 s'explique principalement par le manque de contrôle lors de l'utilisation du RO par le MN ainsi que les failles de sécurité qui y sont inhérentes.

1.1.2 La sécurité des réseaux

Principales attaques dans les réseaux informatiques mobiles

L'usurpation d'identité est l'une des attaques les plus dommageables dans les réseaux informatiques en permettant à un attaquant d'emprunter l'identité de sa victime avec tous les privilèges qui lui sont associés. Il ne s'agit pas simplement de retransmettre des messages interceptés de sa victime, mais plutôt de forger des messages en utilisant l'identité de sa

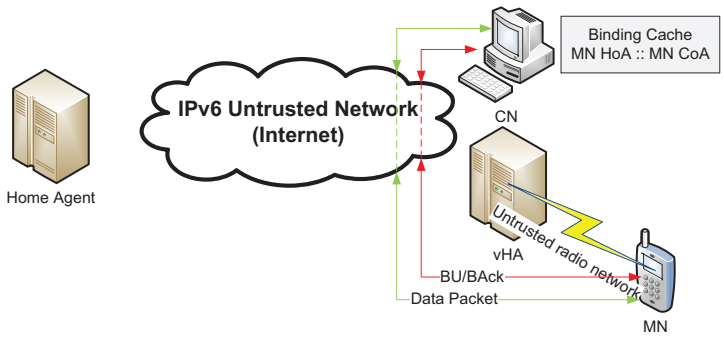


Figure 1.4 Optimisation de route dans MIPv6

victime. Cette attaque est précurseur au man-in-the-middle (MITM) où la communication entre la source et la destination traverse par un intermédiaire malicieux qui peut lire ou modifier le contenu à son gré. Pour qu'elle soit réussie, l'attaquant doit prétendre être la source aux yeux du destinataire et le destinataire aux yeux de la source. Typiquement, l'attaque se limite à usurper uniquement l'identité de l'un des deux rôles pour ainsi lire les messages et les retransmettre au bon destinataire. L'attaquant peut même modifier les messages envoyés par la victime dont l'identité a été usurpée sans que l'autre ne puisse détecter la modification. Le man-in-the-middle est une vulnérabilité commune des protocoles ne pouvant garantir qu'une authentification faible.

Les dénis de service (DoS) représentent une autre catégorie d'attaques très exploitées dans lesquelles l'attaquant épuise assez de ressources de la victime pour qu'il ne puisse plus répondre aux requêtes légitimes. Deux causes sont souvent à l'origine de telles attaques : l'inondation ou une mauvaise implémentation du protocole ou du service. L'inondation est une attaque généralement distribuée (DDoS) par laquelle plusieurs nœuds envoie de multiples requêtes en continue vers sa victime sans se soucier de la réponse. Les requêtes ciblent une opération qui exige un temps de calcul non négligeable et un nombre important de ressources. Elle est extrêmement difficile, voire impossible, à contrer due à sa nature légitime qui n'exploite aucune faille. Le second type d'attaque DoS vise une faiblesse dans l'implémentation des protocoles ou services s'exécutant sur la victime ne nécessitent pas la participation de plusieurs nœuds. En effet, elles s'acharnent sur une vulnérabilité logicielle qui permet à distance de rendre le service (ou encore pire le système en général) instable au point de ne plus pouvoir répondre.

La Figure 1.5 montre d'autres attaques qui sont spécifiquement liées aux MANETs. Ces attaques regroupent celles dites passives, dont le but est de collecter de l'information sans se faire détecter, et actives qui interrompt le bon fonctionnement du réseau en envoyant ou en modifiant des messages. Cette thèse ne porte pas sur les attaques dans les MANETs en soi,

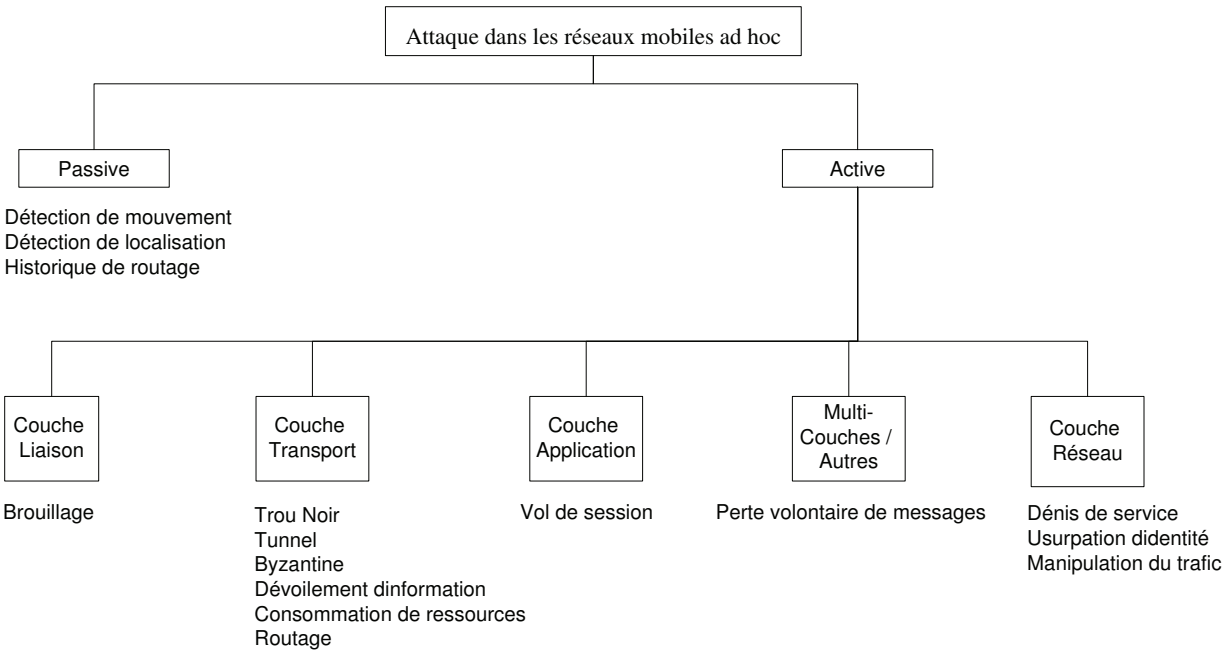


Figure 1.5 Classification des attaques portées dans les MANETs

mais plutôt sur la stimulation des nœuds à collaborer ensemble pour assurer la fonction de routage du réseau.

Objectifs de base d'une solution sécuritaire

L'intégrité, la confidentialité et la disponibilité des données sont les trois objectifs primaires qu'un protocole sécuritaire doit combler. L'intégrité des données signifie qu'aucune altération non autorisée n'est possible et est donc intimement liée à l'authentification. Une authentification faible survient lorsqu'aucune relation n'existe a priori entre les entités impliquées et elles doivent donc s'échanger de l'information pour pouvoir s'authentifier par la suite. Puisque ce premier échange n'est pas authentifié, un attaquant peut facilement s'y introduire en prétendant être une entité légitime. Au contraire, une authentification forte signifie que les parties se connaissent préalablement à un échange authentifié. L'intégrité d'un message l'empêche également d'être rejoué ultérieurement par un attaquant qui l'intercepte et le retransmet dans un autre environnement, indépendamment s'il peut lire ou analyser son contenu.

La confidentialité d'un message assure que son contenu ne peut être lu que par le ou les destinataires. Par conséquent, si une tierce partie intercepte le message, elle ne pourrait le comprendre, l'analyser ou encore y répondre. Cet objectif est crucial lors de l'échange

d'information sensible qui compromettrait le bon fonctionnement du protocole.

Finalement, la disponibilité des données se définit comme étant la capacité de fournir des ressources ou des services à un moment spécifique, ou continuellement pendant un intervalle de temps donné. Les attaques de dénis de service causées soit par une inondation ou l'exploitation d'une faille visent explicitement à contrer cet objectif.

Primitives cryptographiques de sécurité

Les primitives cryptographiques sont des algorithmes de bas niveau utilisés dans les systèmes de sécurité informatique et qui sont en général divulguées publiquement. Les fonctions de hachage (i.e. SHA, MD5, etc.) et de chiffrement (i.e. AES, 3DES, blowfish, etc.) sont les 2 catégories de primitives les plus souvent définies. Une fonction hachage transforme de manière irréversible une entrée en un hash de plus petite taille qui est souvent utilisé comme empreinte cryptographique pour vérifier que l'entrée n'a pas été modifiée. Cependant, puisque tout nœud peut régénérer un nouveau hash, l'empreinte doit être chiffrée pour éviter des modifications non autorisées.

La nature des algorithmes de chiffrement peut être symétrique ou asymétrique tout dépendamment si les clés de chiffrement ont été préalablement partagées ou non. Le chiffrement symétrique est beaucoup plus efficace mais exige une relation préexistante entre les entités pour ne pas devoir divulguer les clés. Ainsi, à la fois l'authentification et la confidentialité des données sont satisfaites si un message est chiffré avec la clé symétrique partagée et secrète. Au contraire, le chiffrement asymétrique définit une paire de clés par entité dont l'une est publique et divulguée, alors que l'autre est privée et n'est connu que par celle l'ayant générée. Ainsi, un message chiffré avec la clé publique d'un nœud ne peut être déchiffré que par la clé privée correspondante, assurant ainsi la confidentialité. D'autre part, si une empreinte cryptographique est chiffrée avec la clé privé, elle peut être validée par la clé publique qui est accessible à tous les nœuds, assurant l'authentification (et l'intégrité par l'utilisation du hash). L'avantage des deux types de clé peut être combiné avec le chiffrement hybride qui permet à deux entités inconnues de s'échanger une clé symétrique en la chiffrant par une clé asymétrique. Le chiffrement hybride est le mécanisme le plus répandu à travers les solutions de sécurité.

La suite de protocoles IPSEC, proposant les entêtes d'authentification (AH) et de chiffrement de données (ESP), a introduit le concept d'association de sécurité (SA) dans laquelle les primitives cryptographiques de sécurité négociées entre deux entités sont regroupées. Chaque SA n'est valide que pour un type de trafic précis et pour une seule direction. Ainsi, un tunnel sécurisé entre deux nœuds nécessite au minimum 2 SAs. Des extensions permettant à un SA de lier logiquement plusieurs nœuds existent et sont particulièrement utilisées dans les

protocoles de diffusion.

1.1.3 La théorie des jeux dans les réseaux informatiques

Les outils de la théorie des jeux servent à prédire la performance d'un réseau ou encore à concevoir des mécanismes incitatifs menant à la réalisation des objectifs définis dans le contexte où les nœuds qui y participent agissent de manière rationnelle. Plus précisément, la théorie des jeux sert à modéliser l'interaction entre 2 ou plusieurs joueurs à comportement rationnel dont la solution est trouvée dans une optimisation à plusieurs niveaux qui tient compte des actions (stratégies) des compétiteurs. Elle s'applique notamment bien dans les MANETs puisque les nœuds qui y participent sont autonomes et de nature décentralisée, donc propice aux comportements égoïstes menant à la satisfaction des intérêts personnels plutôt que collectifs.

Fonction d'utilité et équilibre de Nash

Les préférences des joueurs sont exprimées à travers la fonction d'utilité qui est partagée par tous les joueurs. Quoique les variables qui la composent ont des valeurs différentes d'un joueur à l'autre, l'expression de la fonction d'utilité doit être la même pour tous les joueurs. Elle définit le gain ou la perte d'un joueur suite à une action qu'il a entreprise et considérant les actions possibles des autres joueurs. Dans le cas d'une compétition basée sur le prix, la fonction de profit et de coût sont alors utilisées pour exprimer les préférences des joueurs.

Le point d'équilibre dans un jeu est atteint lorsque tous les joueurs optimisent leur fonction d'utilité selon l'information sur les stratégies des autres compétiteurs qui leur est accessible. Ainsi, lorsque la dérivée de la fonction d'utilité est nulle pour tous les joueurs, représentant du même coup la meilleure action possible face aux actions possibles des autres joueurs, l'équilibre de Nash est atteint. Si pour tous les mouvements possibles des autres, l'équilibre mène toujours vers les mêmes actions pour un joueur donné, l'ensemble de toutes ces actions constitue une stratégie pure. Il peut y avoir plusieurs points d'équilibre de Nash menant à une sélection aléatoire de stratégies pures, nommée stratégies mixtes. Puisqu'une stratégie mixte assigne une probabilité à chaque stratégie pure, elle peut être vue comme une généralisation de cette dernière.

Les types de jeux

Un jeu peut prendre plusieurs formes dépendamment de son contexte et des suppositions posées. La liste ci-dessous décrit différentes caractéristiques d'un jeu (un jeu peut en posséder une ou plusieurs) :

– coopératif ou non-coopératif : un jeu est coopératif lorsque les joueurs forment une coalition pour s'engager dans un même but. La communication inter-joueur devient un élément clé pour un jeu coopératif, alors qu'il est généralement interdit dans un jeu non-coopératif ;
– symétrique ou asymétrique : un jeu est symétrique lorsque les gains ou les pertes pour une stratégie particulière ne dépend que des actions jouées par les autres plutôt que de l'identité des autres joueurs. Ainsi, si pour une même action exécutée par les joueurs, le premier joueur tire un gain différent du joueur 2, le jeu est asymétrique. La Figure 1.6 montrent un exemple de jeu symétrique et asymétrique ;

Figure 1.6 Exemple d'un jeu symétrique (gauche) et asymétrique (droite)

	A	B
A	5,5	0,10
B	10,0	3,3

	A	B
A	5,3	0,10
B	10,0	3,5

– somme nulle ou somme non-nulle : un jeu possède une somme nulle lors que la somme des gains et des pertes de tous les joueurs et pour toutes les combinaisons de stratégies possibles est nulle. En d'autres termes, un joueur ne tire un gain que si un autre joueur subit une perte de même valeur ;
– simultané ou séquentiel : un jeu est simultané si les joueurs ignorent les actions choisies précédemment par les autres. Au contraire, si un joueur considère les stratégies des autres qui ont été jouées précédemment dans sa prise de décision, le jeu est alors séquentiel et le temps devient un facteur important à considérer dans la fonction d'utilité ;
– information parfaite ou imparfaite : si tous les joueurs connaissent les actions sélectionnées par les autres depuis le début du jeu, alors le jeu est basé sur de l'information parfaite. Ce type d'information nécessite au préalable d'avoir un jeu séquentiel puisque par définition, un jeu simultané implique que pas tous les joueurs connaissent les mouvements précédents des autres.

Marché oligopolistique : la compétition de Bertrand vs Cournot

Les oligopoles sont caractérisés un nombre restreint de firmes (vendeurs) face à une multitude de consommateurs (demandeurs). La compétition peut alors s'effectuée sur une base de prix ou de quantité selon les modèles respectifs de Bertrand et de Cournot. Les deux reposent essentiellement sur les mêmes suppositions :

1. Il y a au moins deux firmes qui produisent des produits homogènes ;

2. Les firmes ne coopèrent pas ensemble (aucun cartel toléré) ;
3. Les firmes sont rationnels et agissent pour maximiser leur profit selon la décision des compétiteurs ;
4. Les firmes agissent simultanément en choisissant une quantité (Cournot) ou un prix (Bertrand) ;
5. Les consommateurs achètent en minimisant leurs coûts d'achat.

Le choix du modèle dépend strictement du contexte d'application. Généralement, lorsque la capacité de production peut être facilement modifiée, la compétition de Bertrand s'applique mieux. Il est important d'effectuer le bon choix puisque les implications des deux modèles diffèrent. En effet, dans une compétition de Bertrand, il ne suffit que la présence de deux firmes (dans contexte de joueurs symétriques ayant accès à de l'information parfaite) pour obtenir une compétition parfaite dans laquelle les prix sont réduits au coût marginal des firmes. Ce même résultat n'est atteint par la compétition de Cournot que si le nombre de firmes en compétition est infini. Dans le cadre de cette thèse, la compétition de Bertrand cadre mieux le contexte d'application où la source (consommateur) définit le produit désiré et les firmes entrent en compétition pour lui offrir le meilleur prix.

1.2 Éléments de la problématique

Se pencher sur la sécurité dans les réseaux mobiles de nouvelle génération signifie de considérer à la fois les MANETs, les réseaux d'accès 4G cellulaires ainsi que les applications vidéo. Une approche proactive dans la sécurisation des MANETs en offrant des solutions cryptographiques est problématique d'une part dû à la gestion des clés dans un environnement décentralisé et d'autre part par des ressources très limitées qui poussent les nœuds aux comportements égoïstes. En effet, malgré la disponibilité de telles solutions, encore faut-il que les participants trouvent l'intérêt de les utiliser et de collaborer ensemble afin de satisfaire de manière sécuritaire les fonctions de routage. La mise en place d'incitatifs punitifs avec un système de détection d'intrusion basé sur la réputation des nœuds ou, au contraire, d'incitatifs positifs avec un mécanisme de récompenses ou de compensation devient alors cruciale. Quoique leur but soit commun, les deux écoles de pensée ont des caractéristiques qui divergent. D'une part, les IDS basés sur la réputation des nœuds souffrent d'abord de l'abus des seuils de classe qui permet à un nœud de volontairement se comporter de manière égoïste juste avant d'atteindre le seuil de tolérance maximale qui l'isolera du réseau pour ensuite regagner la réputation perdue en agissant selon les normes. Répété continuellement par plusieurs nœuds, cet abus peut facilement compromettre la stabilité du réseau. De plus, lorsque les nœuds surveillants sont complices aux nœuds malveillants, il sera alors impossible

pour les autres nœuds de détecter les défaillances et d'isoler les coupables du réseau. La Figure 1.7 illustre une attaque dans laquelle la complicité des nœuds malicieux B et C permet une modification non-autorisée du message. Il est à noter que l'instabilité des liaisons sans fil et l'individualité des nœuds présentent des problèmes inhérents à l'évaluation des nœuds voisins et distants pour leur attribuer une réputation précise.

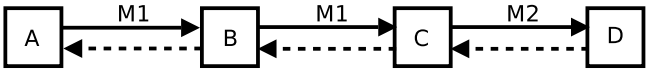


Figure 1.7 Exemple d'une attaque complice entre les nœuds B et C

Les modèles économiques, dans lesquels la monnaie virtuelle, connue sous le nom de nuglets, permet à la source de compenser les nœuds intermédiaires pour les ressources hypothéquées dans la retransmission de ses messages, représentent une alternative intéressante à la réputation qui s'adapte mieux aux caractéristiques des MANETs. Néanmoins, quoique les modèles économiques pour les protocoles de routage dans les réseaux informatiques classiques aient été extensivement étudiés dans la littérature, la réalité des MANETs est trop différente pour qu'ils puissent être repris ou même servir de base. Mise à part la nécessité d'un matériel infraudable, tel qu'une carte à puce, dans chacun des nœuds du réseau pour assurer la bonne gestion des nuglets, les interférences et l'absence d'unité centralisée pouvant favoriser l'échange d'information inter-nœuds viennent considérablement complexifier le modèle notamment lors de l'intégration du support à la QoS. Les solutions existantes se basent sur des suppositions irréalistes pour les MANETs ou encore celles supportant la QoS ne se penchent que sur un critère particulier de QoS très spécifique sans pouvoir servir de base pour étendre le modèle en y ajoutant des critères QoS supplémentaires.

La multidiffusion a été un sujet très convoité dans la dernière décennie par les chercheurs en quête de réduire la bande passante dans les réseaux. L'enjeu est d'autant plus important aujourd'hui avec l'apprivoisement de services vidéo (tels qu'IPTV) de plus en plus gourmands en ressources par les utilisateurs mobiles. Les solutions de sécurisation des protocoles de multidiffusion existantes ne sont pas adaptées à supporter un groupe de nombreux membres dynamiques pouvant s'abonner à plusieurs flots à tout moment et en étant exposés à un environnement à haut taux de perte de paquets. Le besoin d'un protocole de distribution de clés régénératrices n'aura jamais été aussi urgent avec l'arrivée de la convergence des réseaux fixes et mobiles dans les réseaux 4G.

La dernière problématique abordée dans cette thèse concerne le mécanisme d'optimisation de route pour MIPv6. L'absence d'information pré-partagée entre le MN et le CN fait de la sécurisation dans RO un défi de taille. La solution du *return routability* décrite dans

la spécification de MIPv6 exige plusieurs échanges entre le MN, son HA et le CN et n'assure que l'accessibilité du MN à travers ses adresses HoA et CoA plutôt que de sécuriser la communication. En effet, cette solution est vulnérable à l'usurpation d'identité menant vers de multiples attaques très néfastes telles que le man-in-the-middle, redirection des flots et les dénis de services. D'autres solutions lient le préfix de l'adresse du sous-réseau mère à un certificat afin d'assurer l'authentification des MNs à travers son HA, mais leurs suppositions ne sont pas réalistes. En effet, le certificate-based binding update (CBU) nécessite une infrastructure d'authentification fragmentée à travers tous les domaines qui forcerait les opérateurs en compétition à collaborer ensemble en s'échangeant les certificats liés aux préfixes de leurs sous-domaines. Pour pallier la gestion des certificats qui peut être très problématique en considérant plusieurs préfixes de sous-domaine appartenant à un nombre grandissant d'opérateurs, une approche hiérarchique a également été proposée. HCBU se fie sur le déploiement global d'une certification en chaîne à 3 niveaux lié au préfix d'une adresse sous-réseau IPv6 qui permettrait à un nœud voulant valider l'adresse de remonter la hiérarchie de certificats jusqu'à la racine (entité de confiance). Quoique HCBU élimine l'exigence pour les opérateurs de maintenir une infrastructure d'authentification, le consortium IETF n'envisage pas un déploiement 3-tiers de ce genre.

Une approche plus décentralisée à l'authentification est offerte par la génération d'adresses cryptographiques qui permet au MN de générer son adresse à partir de son propre certificat qu'il a lui-même généré. L'usurpation d'identité ne devient alors possible que si l'algorithme de génération d'adresse mène vers l'adresse de sa victime (collision d'adresse) à partir d'un certificat généré. Avec l'emploi du mécanisme d'extension de hachage, le risque de collision est faible. Cependant, CGA est vulnérable à l'attaque du temps-mémoire permettant de mettre en place une table avec toutes les différentes valeurs de Hash-2 et ainsi garantir d'usurper l'adresse de sa victime en un maximum de 2^{59} opérations de hachage. De plus, puisque l'empreinte cryptographique n'inclut pas l'adresse du nœud, CGA est aussi vulnérable aux attaques de rejeu.

Malgré ses faiblesses, CGA a été intégré dans plusieurs solutions alternatives au RR et à celles basées sur les certificats gérés par des autorités de certificats. Parmi celles-ci, le RFC4866 de Arkko *et al.* (2007) reprend le RR mais en permettant au MN de générer son HoA via le CGA et ainsi assurer une authentification faible. Cependant, l'absence d'information pré-partagée entre le MN et le CN permet à un attaquant interceptant les messages initiaux et de s'introduire entre les deux (MITM). De plus, puisque le CoA n'est pas un CGA, tout MN avec une adresse HoA légitime peut usurper l'adresse d'une victime et ainsi lui rediriger ses flots (attaque par redirection). La solution décrite dans le RFC4866 détaille bien ce risque et propose de pallier les effets de manière réactive en s'assurant que le MN réponde au CoA

spécifié dans le BU selon en vérifiant périodiquement selon son degré de confiance. Une telle approche réactive ne sera jamais optimale et ouvre la porte aux abus qui peuvent être néfastes si exploités simultanément par plusieurs attaquants ciblant la même victime.

Les éléments de la problématique présentés se résument aux quatre questions suivantes qui mèneront vers les objectifs de recherche et chacun des quatre articles complétant la thèse :

- Est-il possible d'intégrer le risque d'attaque complice dans le calcul du chemin le plus fiable dans les MANETs et d'éviter les abus des seuils de réputation interclasses ?
- En optant pour un incitatif de récompenses plutôt que punitif, est-il possible d'élaborer un modèle économique selon des suppositions réalistes s'adaptant au protocole de routage dans les MANETs dans le but de supporter plusieurs critères de QoS ?
- Est-ce réalisable de concevoir un protocole sécuritaire pour la multidiffusion de flots vidéo en temps réel dans un contexte d'un très grand nombre d'abonnés sous un environnement mobile à très haut taux de perte de paquets ?
- La conception d'un protocole d'optimisation de route pour MIPv6 offrant une authentification forte décentralisée et basée sur une information pré-partagée non confidentielle et facilement gérée est-elle réalisable dans les limites technologiques existantes présentement dans Internet ?

1.3 Objectifs de recherche

L'objectif principal de cette thèse est de concevoir des solutions sécuritaires pour les réseaux mobiles de nouvelle génération pour favoriser l'apprivoisement de services vidéo mobiles. Ce but nécessite d'une part la coopération entre les nœuds dans les MANETs pour qu'ils retransmettent les messages reçus sans intention malsaine individuelle ou collective. D'autre part, les besoins en ressources des applications vidéo étant sans cessent grandissante et l'industrie des télécommunications qui mise de plus en plus sur le virage vert, l'efficacité énergétique et la réduction de la bande passante sont deux éléments clés de compétitivité. Le second volet se concentre donc sur la sécurisation du protocole de distribution de clés pour la multidiffusion et du mécanisme d'optimisation de route de MIPv6. Plus précisément, les sous-objectifs suivants sont visés :

1. Concevoir un IDS avec une classification plus exhaustive réduisant l'abus des seuils interclasses en plus de considérer le risque d'attaques complices dans le calcul de la fiabilité des chemins ;
2. Comparer cette solution avec le protocole de routage DSR et le IDS original watchdog/pathrater ;
3. Concevoir un cadre de travail modélisant mathématiquement un système économique,

basé sur l'information imparfaite liée à l'incertitude du nombre de compétiteurs et de leurs coûts de production, permettant la fixation de prix face à une requête de route spécifiant des critères de QoS et un prix de réserve ;

4. Trouver l'équilibre de Nash en stratégies mixtes et évaluer l'impact des prix et la performance du réseau face à l'incertitude ;
5. Concevoir un protocole de distribution de clés régénératrices pour la multidiffusion sécuritaire supportant un grand nombre d'abonnés dans un environnement mobile à haut taux de perte de paquets ;
6. Comparer la performance du protocole de distribution de clés régénératrices pour la multidiffusion avec LKH ;
7. Concevoir une version améliorée du CGA offrant une authentification intégrée efficace tout en éliminant le compromis temps-mémoire qui réduit la complexité de l'attaque d'usurpation ;
8. Concevoir une solution sécuritaire d'optimisation de route pour MIPv6 en utilisant des infrastructures existantes ou approuvées par l'IETF assurant une authentification forte et ainsi éliminer toutes vulnérabilités tout en limitant le nombre de messages échangés dans la liaison radio ;
9. Vérifier et valider formellement la sécurité du CGA et du mécanisme d'optimisation de route pour MIPv6 proposés par leur implémentation dans un vérificateur de modèle à travers un vérificateur de protocoles cryptographiques dans le modèle formel.

1.4 Esquisse méthodologique

Les objectifs fixés sont principalement axés sur 4 aspects distincts dont le fil conducteur demeure d'améliorer la sécurité des réseaux mobiles de prochaine génération. Le premier sujet portant sur un IDS dans les MANETs nécessite d'abord la conception d'une classification beaucoup plus exhaustive des nœuds limitant les transitions entre chacune des classes. Ensuite, le nombre de nœuds observateurs ainsi que leur réputation viennent s'ajouter dans le calcul du chemin le plus fiable afin de réduire le risque d'attaques complices qui ne peuvent être détectées. Le IDS proposé ainsi que les composants watchdog et pathrater originaux seront ensuite implémentés dans le simulateur Qualnet sous le language C/C++ afin de les comparer dans diverses configurations de mobilité et du ratio d'attaquants malicieux et complices.

Les sous-objectifs 3 et 4 nécessitent une bonne maîtrise des concepts de la théorie des jeux dans le but de modéliser mathématiquement le protocole de routage selon les suppositions

posées. Après avoir élaboré sur les étapes du protocole en détaillant le contenu des messages et validé le réalisme des suppositions posées, le choix du type de jeu doit être soigneusement sélectionné pour ensuite établir la fonction d'utilité (de profit) selon le type d'information retenu par les joueurs. L'existence de l'équilibre de Nash doit ensuite être prouvée et trouvée. Le comportement des joueurs rationnels doit ensuite être évalué en implémentant le modèle sous MATLAB qui permet la résolution de système d'équations différentielles. L'impact de l'incertitude sur le prix et le respect des critères de QoS durant la durée complète du contrat ne sont que quelques métriques de performance importantes à évaluer.

Vient ensuite le projet sur la multidiffusion dans lequel la revue de littérature identifiera les solutions les plus prometteuses dans un environnement mobile à haut taux de perte de paquets. Il s'agira ensuite de voir comment une chaîne de hachage bidirectionnelle pourra être intégrée au LKH afin d'obtenir des clés régénératrices. La solution proposée est ensuite implémentée dans Qualnet afin de comparer le taux de données indéchiffrables à celui de LKH dans divers scénarios.

Finalement, une compréhension de l'algorithme CGA est nécessaire pour atteindre le sous-objectif 7 et ainsi ajouter une composante aléatoire dans le calcul du hash-2 afin de contrer l'attaque temps-mémoire. La conception d'un protocole d'optimisation de route, impliquant à la fois le HA du réseau mère et celui du réseau visité dans des domaines de confiance, intègrera l'algorithme CGA amélioré dans l'objectif d'offrir une authentification forte. La solution sera implémenté sous AVISPA, un vérificateur de modèles formels à la volée pour s'assurer de n'avoir aucune vulnérabilité.

1.5 Principales contributions de la thèse

Chacun des sujets traités dans cette thèse ont mené à des contributions, certaines plus importantes que d'autres. D'abord, la recherche effectuée dans les IDS des MANETs a permis notamment d'utiliser le positionnement des nœuds ainsi que la réputation des nœuds surveillants comme facteurs influençant le risque de collusion entre deux nœuds malicieux consécutifs. Quoique mineure, cette contribution ouvre une nouvelle voie de recherche en favorisant l'exploitation de toutes les informations disponibles sur le réseau, plutôt que de se limiter uniquement à la réputation des nœuds intermédiaires pour caculer la fiabilité d'une route.

Une seconde contribution plus innovatrice vient de la conception d'un modèle économique basé sur la compétition de Bertrand pour stimuler les nœuds à coopérer pour satisfaire les critères de QoS durant toute la durée du contrat. Ce modèle est le premier à inclure une incertitude de participation de la compétition qui est fonction du prix fixé et du coût de

production des compétiteurs tout en considérant leur nature asymétrique. Se limitant uniquement à la bande passante, ce cadre théorique ne forme qu'une base qui peut être étendue dans le but d'y intégrer d'autres critères de QoS. L'avantage principal vient des suppositions posées qui sont réalistes dans le contexte des réseaux mobiles ad hoc, une critique présente dans la plupart des modèles appliqués dans le domaine.

À un niveau moins théorique, le protocole de distribution de clés pour la multidiffusion proposé est le seul à pouvoir supporter à la fois les clés régénératrice et un très grand bassin d'abonnés sans voir une dégradation importante de la performance avec le temps. Malgré sa sensibilité aux fréquentes révocations d'usagers, cette solution vient à point dans un contexte de réseau de prochaine génération dont un des buts principaux est d'offrir des services vidéo (IPTV) en permettant une relève verticale sans interruption entre divers réseaux fixes et mobiles caractérisés par un taux non-négligeable de perte de messages. Les clés régénératrices, permettant à un abonné de continuer à pouvoir déchiffrer les données malgré avoir manqué un certain nombre de clés consécutifs, n'étaient disponibles que pour SDR. Cependant, la performance de SDR se dégrade rapidement avec le temps plus le nombre d'usagers s'abonnent au flot de multidiffusion. LKH au contraire offre une performance beaucoup plus stable, mais ne supporte pas les clés régénératrices. Cette solution vient donc pallier des problèmes importants limitant le développement technologique au niveau des applications sensibles à la perte de paquets.

Pour terminer, deux autres innovations intéressantes sont issues suite à la recherche portée sur la sécurisation de l'optimisation de route pour MIPv6. D'une part, l'intégration d'une chaîne de hachage inversée dans le calcul du hash-2 permet à la fois une authentification faible et de contrer les attaques temps-mémoire. De plus, l'algorithme d'extension de hachage peut s'exécuter en arrière-plan et ainsi utiliser la prochaine clé de la chaîne de hachage inversée lorsque le nœud est invité à générer une nouvelle adresse (comme dans le cas d'une relève). La version améliorée du CGA propose également de lier plusieurs adresses ensemble lorsqu'un lien logique existe entre elles permettant une certaine hiérarchie où une entité d'autorité et de confiance autorise l'utilisation de l'adresse générée. C'est notamment le cas dans MIPv6 où le MN possède le HoA et le CoA qui sont respectivement liés au HA du réseau mère et visité. Ainsi, la solution d'optimisation de route sécuritaire proposée innove d'une part en liant le HoA avec le HA, et le CoA avec le vHA. Le MN doit donc faire autoriser ses adresses aux autorités de confiance du domaine (HA et vHA) qui leur enregistre et associe un sous-domaine unique. D'autre part, la solution se base sur les domaines de confiance comme source suffisante d'information pour assurer une authentification forte. Avec la fin du déploiement global de DNSSEC prévu pour la fin 2011, la solution se base sur des suppositions réalistes exigeant une simple gestion de noms de domaines de confiance pour permettre à deux entités

qui a priori ne se connaissent pas de s'authentifier à travers les empreintes cryptographiques des entités de confiance.

1.6 Plan de la thèse

Suivant l'introduction, le chapitre 2 présente une revue critique de la littérature sur les quatre sujets abordés dans cette thèse. Le prochain chapitre porte sur la conception d'un IDS qui intègre le risque d'attaques complices dans le calcul des chemins et est détaillé dans l'article intitulé *Collusion-resistant reputation-based intrusion detection system for MANETs* publié dans *International Journal of Computer Science and Network Security (IJCSNS).*

Toujours dans les MANETs, le chapitre 4 présente un article intitulé *An extensible game-theoretic framework based on Bertrand competition for QoS support in MANETs with uncertain asymmetric costs* qui a été soumis dans la revue *IEEE Transactions on Mobile Computing.* Ce dernier utilise les outils de la théorie des jeux pour proposer un modèle économique basé sur la compétition de Bertrand afin de stimuler la coopération entre les nœuds d'une même route à satisfaire les exigences de QoS pendant une durée définie dans un contexte où les joueurs sont asymétriques et accèdent à de l'information imparfaite sur les compétiteurs.

Le protocole de distribution de clés régénératrices pour la multidiffusion intitulé *An efficient and secure self-healing scheme for LKH* est présenté au chapitre 5. L'article correspondant a été publié dans la 3e édition spéciale *Security and Management* de la revue *Journal of Network and Systems Management* de l'éditeur *Springer.*

Le chapitre 6 présente le dernier article intitulé *Secure Route Optimization for MIPv6 using Enhanced CGA and DNSSEC* qui a été accepté pour publication dans l'édition spéciale *Mobility Management and Security Issues for Mobile Wireless Networks* de la revue *Journal of Telecommunications Management* de l'éditeur *Henry Stewart publications.* La version améliorée du CGA et le mécanisme sécuritaire de l'optimisation de route pour MIPv6 y sont détaillés.

La discussion générale en regard des aspects méthodologiques et des résultats en lien avec la revue critique de littérature suit au chapitre 7.

Pour terminer, le chapitre 8 conclut la thèse en synthétisant les travaux qui ont été effectués et en exposant les limitations sur chacun des sujets abordés.

CHAPITRE 2

REVUE CRITIQUE DE LA LITTÉRATURE

Ce chapitre se consacre à une revue critique de la littérature pour les quatre sous-problèmes considérés dans la sécurisation des réseaux mobiles de nouvelle génération. D'abord, les plus importants systèmes de détection d'intrusion basés sur la réputation des nœuds dans les MANETs sont détaillés et leurs vulnérabilités sont décrites. En second, les modèles économiques pour fixer une somme incitant les nœuds intermédiaires à utiliser leurs ressources pour transférer les messages d'une source sont étudiés. Pour le troisième sous-problème, les protocoles de distribution de clés pour la multidiffusion sont présentés en explicitant leurs domaines d'application suggérés. Pour terminer, les algorithmes de génération d'adresses cryptographiques CGA et CGA++ sont détaillés et les différents types de solutions pour la fonction d'optimisation de route de MIPv6 sont présentés.

2.1 Les incitatifs à la coopération inter-nœuds dans les MANETs

Les nœuds dans les MANETs possèdent des ressources très limitées qui les poussent qu'à satisfaire leurs propres intérêts. Sans aucun incitatif, il n'y a donc aucune raison de croire qu'un nœud consommera volontairement une partie de ses ressources pour transférer le message d'un autre nœud. Or, pour obtenir un réseau fonctionnel sans infrastructure central autoritaire, les nœuds doivent combler les fonctions de routage. Les incitatifs peuvent être de nature punitive ou bénéfique en utilisant respectivement la réputation des nœuds ou encore la monnaie virtuelle comme moyens pour stimuler les nœuds.

2.1.1 Les systèmes de détection d'intrusion basés sur la réputation des nœuds pour les MANETs

La réputation d'un nœud est une évaluation subjective basée sur le comportement de celui-ci qui peut être partagée avec les voisins. Les pionniers (Marti *et al.* (2000)) ont distingué deux composants distincts formant le IDS : le watchdog et le pathrater. Le watchdog s'occupe d'évaluer et de classifier les nœuds découverts selon les observations effectuées localement ou encore reçues par d'autres nœuds observateurs (Michiardi et Molva (2002); Buchegger et Boudec (2002)). Une modification non-autorisée au messsage originale ou encore la non-retransmission du message incrémente le compteur d'actions malicieuses du nœud évalué et s'il atteint un certain seuil, le nœud est alors isolé temporairement du réseau ou encore de

manière permanente si le nombre de récidives est trop élevé. Le pathrater calcule la fiabilité d'un chemin selon la réputation des nœuds intermédiaires et choisit celui ayant la plus haute moyenne, contrairement à la vaste majorité des protocoles de routage pour les MANETs qui aurait sélectionné le plus court chemin.

Dans la solution originale (Marti *et al.* (2000)), les observations se limitent simplement à s'assurer que le message a été retransmis sans modification. Ainsi, un observateur, qui peut également faire parti de la route, sauvegarde temporairement le message original retransmis (ligne pleine) et le compare à la réponse passive (ligne tiretée) reçue du nœud évalué (Figure 2.1). Il est important de noter que les particularités des liens sans fil, comme l'inter-

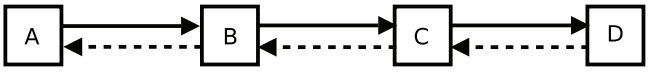


Figure 2.1 Exemple d'observations par réponses passives

férence, qui sont hors de contrôle des nœuds peuvent causer la détection d'un faux positif. Pour cette raison, les nœuds malicieux ne sont jamais isolés de manière permanente. Les simulations montrent que les nœuds malicieux répétant les mêmes actions continuellement sont détectés efficacement. Cependant, il suffit aux attaquants de jauger leurs actions selon les seuils de réputation des différentes classes ou encore de profiter des nœuds qui ne l'ont pas encore évalué pour continuer à opérer sans entrave.

Les différents IDS pour les MANETs se distinguent principalement par le type d'observations effectuées sur les nœuds évalués et les classes qui y sont associées, ou encore si les observations indirectes reçues de voisins distants sont admises. CONFIDANT (Buchegger et Boudec (2002)), par exemple, permet les alertes ciblant les actions malicieuses des nœuds distant et affectant négativement leur réputation, de circuler à travers le réseau. Cette stratégie est inévitablement vulnérable aux attaques de diffamation. CORE (Michiardi et Molva (2002)) est similaire à CONFIDANT mais ne permet que la propagation d'alertes par renforcement positif dans lesquelles les nœuds sont félicités par leurs bonnes actions. Uniquement les observations directes peuvent diminuer la réputation d'un nœud.

La Table 2.1 compare les IDS les plus connus selon différents critères. Cependant, peu importe les types de comportements observés ou les échanges d'alertes entre les nœuds, aucun IDS ne considère le risque d'attaques complices dans lesquelles le message peut être modifié sans que la source en soi averti et que la réputation du nœud fautif ne descende.

Tableau 2.1 Comparaison du Watchdog/Pathrater, CONFIDANT et CORE

		Watchdog / Pathrater	CONFIDANT	CORE
Acceptation des alertes à distance	observations négatives	Non	Oui	Non
	observations positives	Non	Non	Oui
Détection de comportements non-conformes	Modification de messages de routage	Non	Oui	Non
	Modification des données	Oui	Oui	Non
	Non-retransmission des messages de routage	Non	Oui	Oui
	Non-retransmission des données	Oui	Oui	Oui
Isolement permanent des nœuds		Non	Oui	Oui

2.1.2 Les modèles économiques basés sur la théorie des jeux

La monnaie virtuelle (nommé nuglets par Zhong *et al.* (2003)) a mené vers la conception de modèles économiques régissant la fixation de prix pour des biens et services par les joueurs rationnels. Contrairement à la réputation qui nécessite une surveillance constante des nœuds voisins et qui est sensible aux caractéristiques intrinsèques des liens sans fil qui peuvent mener à l'isolement de nœuds légitimes (faux positifs), la monnaie ne sert que de moyen de compensation aux nœuds intermédiaires pour le service de retransmission rendu. Elle suppose cependant la gestion sécurisée de nuglets qui peut être comblée par l'entremise d'une autorité centralisée ou décentralisée telle que par l'utilisation des cartes à puces.

Les équilibres de la compétition de Bertrand

Les équilibres de Nash des modèles basés sur la compétition de Bertrand ont été étudiés sous diverses caractéristiques de jeu et de suppositions. L'exemple typique est le paradoxe de Bertrand Tissier (1984) dans lequel deux firmes symétriques entrent en compétition en fixant des prix simultanément en connaissant le prix maximal que les consommateurs sont prêts à payer. L'unique équilibre de Nash mène vers une compétition parfaite où les prix égalisent le coût marginal de production, et donc vers un profit nul. Dans le cas où la fonction de coût est strictement convexe, indiquant un coût unitaire grandissant avec le nombre d'unités

produits, les auteurs Dastidar (1995) et Hoernig (2002) prouvent l'existence d'un équilibre en stratégies mixtes.

Or, notamment dans les MANETs, il est rare, voire impossible, de s'assurer de l'homogénéité des nœuds. Il est donc plus approprié de supposer l'asymétrie des firmes. Dans un contexte de coûts marginaux constants, Blume (2003) montre que l'équilibre dans un duopole est atteinte lorsque la firme avec le coût de production le plus bas fixe un prix égale au coût de production de la seconde firme. La même conclusion règne en considérant une oligopole à plusieurs joueurs puisqu'uniquement les deux firmes avec les coûts de production les plus bas participent à la compétition (Marquez (1997)).

Les modèles considérés supposent encore l'accès à de l'information parfaite sur les coûts de production des compétiteurs. Dans un contexte de jeu non-coopératif, la communication entre les joueurs est inexistante et force donc les joueurs à estimer les coûts des compétiteurs selon une distribution probabiliste. Les auteurs Vives (1999) et Spulber (1995) montrent l'existence d'un équilibre bayésien de Nash dans un modèle qui associe la perception des coûts de production des compétiteurs à une fonction de densité de probabilité continue à support compact lorsque les coûts marginaux demeurent constants.

Parallèlement à l'estimation es coûts des compétiteurs, leur participation à la compétition est elle aussi incertaine. Janssen et Rasmusen (2002) ont intégré une probabilité d'activité constante dans le modèle de compétition de Bertrand avec firmes symétriques pour trouver l'existence d'un équilibre à stratégies mixtes où les profits de l'industrie sont positifs et descendent avec le nombre potentiel de firmes entrant dans le marché. Ainsi, la compétition parfaite n'est atteinte que dans le cas d'un nombre infiniment grand de firmes potentiels, et une compétition moins féroce fait monter les prix jusqu'à atteindre le prix monopolistique. Il faut cependant noter que ce modèle suppose une probabilité d'activité qui est constante et qui ne dépend pas de la perception des coûts de production des compétiteurs. Or, une firme qui croît posséder un coût de production moindre à celui de ces compétiteurs va nécessairement avoir tendance à participer, alors qu'au contraire, une firme qui croît être désavantagé risque de ne pas y participer. La probabilité d'activité devrait donc être fonction de la perception des coûts de production des compétiteurs. En second lieu, le modèle proposé suppose des firmes symétriques qui partagent le même coût marginal.

L'application de la théorie des jeux dans les réseaux informatiques mobiles

L'usage de la théorie des jeux dans les réseaux informatiques classiques n'est pas récent, comme en témoigne Dasilva *et al.* (2000) et Lazar *et al.* (1997). Ce n'est que depuis quelques années que l'intérêt de l'appliquer dans les réseaux mobiles, notamment les MANETs, s'est manifesté. Qiu et Marbach (2003) se questionnent sur la bande passante optimale à réserver

que les nœuds intermédiaires doivent sélectionner à la requête de la source et y fixer un prix. Le jeu itératif et non-coopératif proposé se base sur la demande externe et les ressources restantes en bande passante et en énergie pour construire la fonction d'utilité et l'optimiser. Une telle approche présente des simplifications qui rendent le modèle défaillant à plusieurs niveaux. D'abord, l'absence de communication entre les nœuds intermédiaires (joueurs) les force à proposer une bande passante qui diffère l'une de l'autre. Or, puisque la bande passante d'un chemin est limité par la plus petite bande passante allouée, il existe un risque important de sur-allocation pour un ou plusieurs nœuds intermédiaires. Cette sur-allocation pousse à la hausse non seulement les prix, mais également le taux de rejet de requête de service par manque de ressources non-allouées. D'autre part, les auteurs ne considèrent pas les interférences ou autres caractéristiques propres aux MANETs comme la mobilité et la fiabilité des liens radio. À ce propos, les auteurs dans Lu et Pooch (2004) ont repris les solutions de marchandage de Nash en l'étendant à plusieurs joueurs pour conclure qu'une compensation égale doit être donné aux nœuds intermédiaires et aux voisins subissant les interférences.

Les outils de la théorie des jeux permettent également de concevoir un protocole de routage avec des règles de jeu qui mènent vers un équilibre d'optimum social (Pareto) qui respecte les objectifs désirés. COMMIT (Eidenbenz *et al.* (2008)) et Ad Hoc-VCG (Anderegg et Eidenbenz (2003)) sont deux exemples qui couvrent la découverte de route et les phases de retransmissions de messages. Les auteurs montrent que la stratégie optimale des joueurs est de divulguer leurs informations privées réelles et qu'ils n'ont aucun intérêt à mentir. Toutefois, les auteurs ne détaillent aucunement comment les joueurs définissent leurs coûts et fixent leurs prix. De plus, COMMIT suppose que la destination agisse toujours légitimement en choisissant la route qui coûte la moins chère.

De manière globale, toutes les solutions proposées négligent les applications en temps réel qui requièrent des critères de QoS très strictes et qui sont généralement définis par la source et non pas dictés par les nœuds intermédiaires. Finalement, aucune solution ne considère la compétition inter-routes plutôt qu'intra-route. Quoique les nœuds doivent être considérés comme des entités indépendantes et rationnels, durant les phases de découverte de route et de retransmission de messages, les nœuds intermédiaires d'une même route collaborent ensemble et entrent en compétition contre les autres nœuds intermédiaires formant des routes alternatives. Cette distinction importante n'est considérée dans aucune solution existante.

2.2 Les protocoles de distribution de clés pour la multidiffusion

La multidiffusion permet à un groupe d'abonnés de recevoir simultanément un message à l'aide d'une seule transmission du message par la source. La tâche de dupliquer le message aux

bonnes interfaces repose donc sur les routeurs du réseau. La sécurité dans la multidiffusion repose essentiellement sur la confidentialité des clés expirées à l'arrivée de nouveaux abonnés (*backward key secrecy*) et des nouvelles clés à la révocation d'abonnés existants (*forward key secrecy*). La plupart des protocoles de distribution de clés sont hiérarchiques et le serveur de contrôleur de clé de groupe (CGKS) crée et maintient un arbre dont chacune des feuilles est associée à un abonné et chaque racine intermédiaire est un nœud logique possédant une clé partagée avec un ou plusieurs abonnées. Les protocoles Logical Key Hierarchy (LKH) (Wong *et al.* (2000); D. *et al.* (1999)) et Subset-Difference Revocation (SDR) (Naor *et al.* (2001)) sont les deux protocoles de distribution de clé pour la multidiffusion les plus populaires. LKH doit maintenir l'état du système à travers son arbre hiérarchique pour effectuer une mise-à-jour ciblée des clés de chiffrement. Au contraire, SDR est un protocole sans état puisque l'emplacement des abonnés dans l'arbre est invariant et leur permet de recevoir une information initiale nécessaire pour déchiffrer les nouvelles clés de trafic soit directement ou après une manipulation de cette dernière (fonction de hachage).

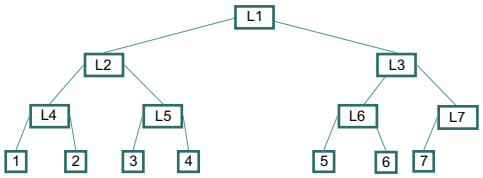


Figure 2.2 Exemple d'arbre binaire hiérarchique

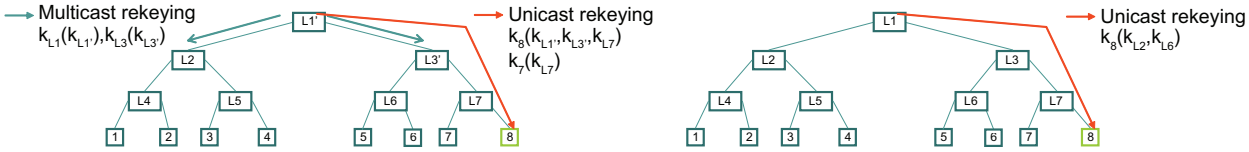
La différence entre LKH et SDR repose essentiellement sur l'information initiale reçue issue de la divergence dans la création et la gestion de l'arbre. Un abonné sous LKH reçoit les clés des nœuds logiques qui se trouvent plus haut que la feuille le représentant dans l'arbre hiérarchique. Sous SDR, les clés envoyés par le GCKS sont celles des nœuds logiques représentant les sous-ensembles dans les branches non-liées à la feuille de l'abonné. De plus, les clés dans les nœuds logiques sous SDR sont généralement liées à l'aide d'une fonction de hachage de gauche (F_G) et de droite (F_D). Soit l'arbre initial montré à la Figure 2.2, K_{LX} la clé du nœud logique LX et K_X la clé privé de l'abonné X, où X est un entier naturel, l'équivalence des clés sous SDR est :

- $K_{L2} = F_G(K_{L1})$;
- $K_{L3} = F_D(K_{L1})$;
- $K_{L4} = F_G(F_G(K_{L1})) = F_G(K_{L2})$;
- $K_{L5} = F_D(F_G(K_{L1})) = F_D(K_{L2})$;

– $K_{L6} = F_G(F_D(K_{L1})) = F_G(K_{L3})$;
– $K_{L7} = F_D(F_D(K_{L1})) = F_D(K_{L3})$;

Puisque l'arbre doit demeurer inchangé en tout temps sous SDR, il est important d'estimer correctement le nombre potentiel d'usagers pour créer un arbre pouvant supporter tous les abonnés.

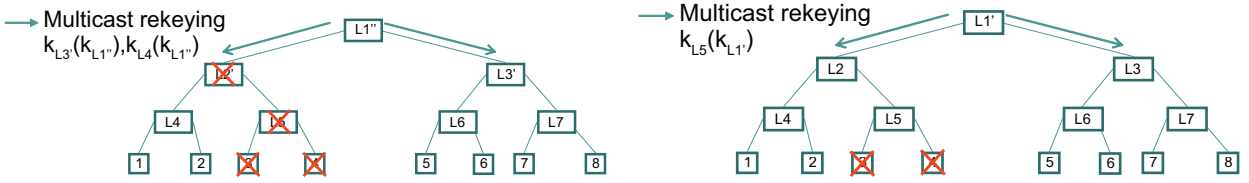
Sous LKH, lorsqu'un abonné se joint au groupe de multidiffusion, il reçoit les clés de chiffrement de tous les nœuds logiques faisant parti de la branche allant de la racine jusqu'à la feuille qui lui correspond. Au contraire, sous SDR, il reçoit les clés des nœuds logiques au sommet des sous-ensembles non-liés à la feuille qui associe l'abonné nouvellement joint. À partir de ces clés, il peut trouver les clés des nœuds logiques de plus bas niveaux en appliquant la bonne fonction de hachage. La Figure 2.3 montre un exemple où un nœud se joint au groupe selon LKH et SDR. De plus, si un nœud quitte temporairement le groupe pour y rejoindre par la suite, SDR replace le nœud à la même feuille qu'il avait quitté et aucun transfert de clé n'est nécessaire puisque l'arbre, et par conséquent les clés, demeure toujours inchangé. Le coût de la mise-à-jour des clés pour une arrivée sous LKH est relativement constant.



(a) Messages échangés dans LKH (b) Messages échangés dans SDR

Figure 2.3 Exemple d'exécution à l'arrivée d'un nouvel abonné pour LKH et SDR

Lors d'une révocation dans LKH, les nouvelles clés sont chiffrées par les clés des nœuds logiques non-liés aux abonnés révoqués. Sous SDR, ces mêmes clés sont chiffrées par les clés des nœuds logiques du sous-ensemble plus haut dans la hiérarchie. La Figure 2.4 montrent les échanges de clés engendrés par la révocation d'un abonné sous LKH et SDR respectivement. Dans l'exemple, les nœuds révoqués appartiennent au même sous-ensemble ce qui rend

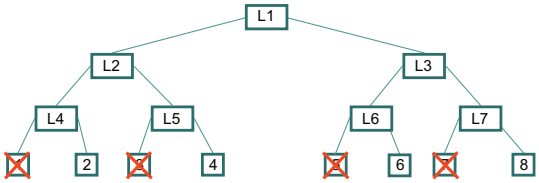


(a) Messages échangés dans LKH (b) Messages échangés dans SDR

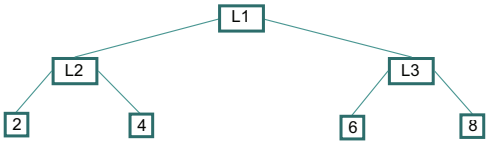
Figure 2.4 Exemple d'exécution au départ de deux abonnés pour LKH et SDR

SDR plus efficace que LKH. Cependant, dans le cas plus général, alors que l'arbre peut être

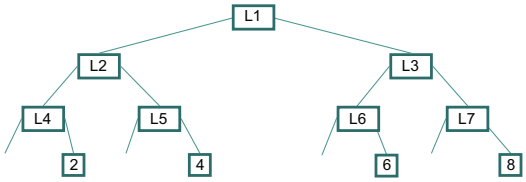
facilement balancé sous LKH, la révocation de plusieurs membres appartenant à différents sous-ensembles augmente drastiquement la complexité des messages de mise-à-jour des clés sous SDR comme montre la Figure 2.5 (Chen et Dondeti (2003a); Zhu *et al.* (2003); Zhu et Jajodia (2003); Dutta *et al.* (2007)).



(a) Exemple de révocation



(b) Arbre résultant sous LKH



(b) Arbre résultant sous SDR

Figure 2.5 Gestion de l'arbre logique sous LKH et SDR

Dans un contexte d'IPTV où le bassin d'abonnés est très grand et le nombre de révocations peut devenir très significatif, la dépendance des événements à la complexité des messages de mise-à-jour des clés de SDR rend LKH un meilleur candidat comme protocole de distribution de clés dans la multidiffusion.

D'autres solutions telles que les arbres à fonction de hachage (OFT) Sherman et McGrew (2003) et ELK Perrig *et al.* (2001) utilisent une approche semblable à LKH et SDR sans apporter un réel avantage.

2.2.1 Distribution fiable de clés

Dans un environnement mobile, les pertes de messages arrivent fréquemment. Or, LKH est extrêmement sensible à la perte d'un message de mise-à-jour des clés qui exige alors aux abonnés ciblés par le message d'effectuer à tour de rôle une requête vers le GCKS pour une retransmission. Les protocoles de distribution fiable de clés visent à diminuer ce risque en optant pour la redondance partielle d'information. Le Proactive FEC (Zhang *et al.* (2003))

utilise le code de Reed-Solomon afin d'envoyer en multidiffusion des bloques d'information redondante (parité) permettant la reconstruction des clés manquantes simultanément par les abonnés ayant subit une perte de messages de mise-à-jour des clés. Il faut cependant bien calibrer la taille des bloques afin d'éviter d'inonder les abonnés qui ont un faible taux de perte de paquets. Le protocole de distribution fiable WKA-BKR (Setia *et al.* (2002)) est une technique simple qui consiste en la retransmission des clés groupés selon la probabilité qu'elles soient mal reçues par les destinataires. En ciblant ainsi les abonnés susceptibles de ne pas recevoir les clés et en leur envoyer les mêmes clés plusieurs fois, il y a une plus grande chance qu'ils les reçoivent. Une méthode hybride, le WFEC-BKR (Zhu *et al.* (2003)), propose de remplacer l'envoi direct des clés dans WKA-BKR par des bloques de parité comme dans Proactive FEC. Il est à noter que les protocoles de distribution fiable imposent une certaine surcharge au réseau et ne sont pas aussi efficaces que les clés régénératrices. Cependant, ils peuvent s'appliquer dans LKH et ainsi supporter un plus grand bassin d'utilisateurs.

2.2.2 Distribution de clés régénératrices

Une alternative plus efficace consiste en la distribution de clés régénératrices permettant ainsi à un abonné de manquer un certain nombre de messages de mise-à-jour des clés consécutifs avant de ne plus pouvoir déchiffrer le trafic. SDR se démarque à ce sujet avec la possibilité de régénérer les clés manquantes (Zhu *et al.* (2003)). Soit m le nombre maximal de clés manquantes consécutives alloué, l'idée est de diviser les abonnés en m+1 groupes selon leur durée d'inscription et de transmettre les m clés précédentes de groupe chiffrées avec la clé courante individuellement en multidiffusion. Afin d'empêcher un nouvel abonné d'accéder aux anciennes clés ou encore de partager les clés avec d'autres abonnés qui se joints à différents moments, la clé de chiffrement courante est lié au temps et à une durée de session. Pour ce faire, la clé de chiffrement courante est associé (à l'aide d'un XOR) à une clé précédente inconnue par le nouvel abonné. Pour la durée de la session, une chaîne de hachage de taille m+1 est construite en appliquant itérativement la fonction de hachage H : $K^m(i), K^{m-1}(i), K^{m-2}(i), ..., K^0(i)$ où $K^0(i) = H(K^1(i)) = H^2(K^2(i))$ est la clé de groupe de la racine que tous les abonnés doivent posséder pour déchiffrer le trafic. Ainsi, un abonné ayant reçu la clé $K^{i-j}(i)$ au temps j peut retrouver les clés $K^{i-j-1}(i), ..., K^0(i) \forall i - m < j < i$ et ainsi déchiffrer un maximum de j clés, d'où vient la propriété régénératrice des clés distribuées. Il est cependant essentiel que l'abonné ait reçu de manière fiable la clé courante pour qu'il puisse par la suite bénéficier des clés régénératrices qui lui permettront de manquer jusqu'à j messages de mise-à-jour des clés consécutifs avant de ne plus pouvoir déchiffrer le trafic.

2.3 Les solutions d'optimisation de route pour MIPv6

Permettre une communication directe entre le nœud mobile (MN) et le nœud correspondant (CN) sauve une bande passante considérable sur le réseau mère du MN qui n'aura plus à servir d'intermédiaire. Par contre, une telle approche enlève tout contrôle à l'opérateur pour s'assurer que le MN opère selon les politiques du réseau mère ou encore pour le facturer selon l'usage. De plus, supposant qu'aucune restriction n'est imposée au MN par son opérateur, encore faut-il que la communication entre le MN et le CN puisse être sécurisée. Ainsi, puisque tous les nœuds mobiles profitent d'une liberté complète sur les messages envoyés et les destinataires, plusieurs attaques peuvent être portées sur différents réseaux ou encore des victimes ciblées. Le but minimal recherché des solutions proposées est donc d'assurer une authentification entre le MN et le CN dans un contexte marqué par l'absence d'information pré-partagée entre deux entités a priori inconnues l'une de l'autre.

2.3.1 La solution standardisée d'optimisation de route du return routability dans MIPv6

Le *return routability (RR)* est la solution de base proposée dans le protocole MIPv6 (Johnson *et al.* (2004)). Pour débuter, le MN envoie deux requêtes au CN contenant chacune un témoin distinct. La première requête nommée *Home Token Init (HoTI)* traverse le réseau mère du MN et provient de l'adresse de réseau mère attribuée au MN (HoA). Quant au *Care-of Init(CoTI)*, la seconde requête provient de l'adresse que le réseau visité a attribuée au MN (CoA) et se rend directement au CN. Suite à la réception des deux requêtes, le CN génère une clé d'authentification à partir d'une clé maintenue secrète par le CN, des deux adresses du MN (HoA et CoA) et de deux *nonces* créés aléatoirement. Chaque moitié de cette clé générée sera envoyée en clair au HoA et au CoA du MN faisant ainsi traverser le *Home Token (HoT)* et le *Care-of Token (CoT)* à travers deux réseaux distincts. Le MN reconstruit ainsi la clé avec une simple concaténation et envoie une mise-à-jour du lien (*binding update (BU)*) authentifiée avec cette clé au CN.

Ce mécanisme consiste principalement en la vérification de l'accessibilité du MN à travers ses deux adresses et ne sécurise aucunement l'optimisation de route dans MIPv6, rendant ainsi le RR vulnérable à de multiples attaques :

1. Détournement de session (*Session Hijacking*)
   Il suffit qu'un attaquant intercepte le témoin mère dans le HoT pour qu'il puisse détourner une communication existante entre un MN et un CN vers lui même. Il débute par forger un CoTI avec sa propre adresse IP (au lieu du CoA du MN victime) et l'envoie ensuite au CN qui lui répondra avec un CoT. L'attaquant concatène alors le HoT

intercepté et le CoT reçu pour authentifier le BU qui remplacera le CoA de sa victime avec l'adresse de l'attaquant qui aura comme conséquence de détourner le flot que le CN avait avec la victime vers l'attaquant. Pour que cette attaque soit réussie, le CoTI doit être reçu par le CN avant la mise-à-jour des *nonces*. Autrement, l'attaquant devra soumettre également un HoTI et intercepter le HoT à nouveau.

2. L'homme du milieu (*Man-in-the-middle*)
   Une version plus transparente de l'attaque du détournement de session est celle de l'homme du milieu dans laquelle l'exécution est similaire à l'exception que l'attaquant retransmet les données reçues du CN vers sa victime. Si le flot n'est pas chiffré, l'attaquant peut le visionner et l'absence de signature lui permettrait de le modifier.
3. Les dénis de service (*Denial of Service (DoS)*)
   Le RR est vulnérable à de multiples attaques de DoS visant le CN. Ayant intercepté N_{HoT} témoins HoT et N_{CoT} témoins CoT du même CN, un attaquant pourrait exiger au CN de rediriger $N_{HoT} \times N_{CoT}$ flots en très peu de temps. Une autre attaque plus simple serait d'envoyer plusieurs requêtes HoTI et CoTIs au CN pour qu'il génère une multitude de témoins en peu de temps. Malgré la facilité d'exécution de ces attaques, une bonne implémentation du RR limitant le nombre d'opérations à exécuter simultanément viendrait corriger cette problématique. Les attaques de DoS dans le RR peuvent également cibler les réseaux visités. Il suffit qu'un attaquant s'y connecte, envoie les requêtes de RR et répond avec un BU authentifié pour rediriger ses flots vers le réseau connecté. En répétant de multiples fois cette opération pour ensuite se déconnecter du réseau visité, il exploite le délai de vérification de l'accessibilité des adresses du MN par le CN (plusieurs minutes) pour inonder le réseau visité.

2.3.2 Les solutions d'optimisation de route basées sur les certificats

Les attaques du détournement de session et de l'homme du milieu sont possibles dans RR dues à l'absence d'authentification. Une solution serait de ne permettre des opérations de RO que par des entités de confiance. Pour ce faire, le Certificate-based Binding Update (CBU) (Deng *et al.* (2002)) exige le déploiement d'une autorité centrale de certificats (CA) pour chacun des opérateurs afin de lier le préfixe du sous-réseau de l'adresse du réseau mère (HLSP) avec un certificat dont la clé privée sera gardée secrètement par le HA. Ainsi, l'établissement du RO est initié par le MN qui envoie une requête à son HA et ce dernier se charge de fournir le certificat lié au HLSP en prouvant son authenticité à l'aide d'une signature. Une clé secrète d'authentification de BU est alors partagée entre le HA et le CN suite à un échange Diffie-Hellman authentifié par le certificat. Le HA retransmet alors cette clé au MN à travers un

tunnel sécurisé préexistant et ce dernier envoie alors un BU authentifié au CN pour finaliser l'exécution.

Malgré l'apport de propriétés de sécurité intéressantes vis-à-vis le RR, cette solution repose sur une infrastructure d'authentification fragmentée à travers les domaines de tous les opérateurs. Une telle supposition est irréaliste et pose de sérieux problèmes d'extensibilité et de flexibilité. En effet, les différentes administrations sont généralement réticentes à partager leurs informations ou à faire confiance à un CA appartenant à l'un de ses compétiteurs. De plus, CBU ne valide pas l'adresse CoA spécifié dans la requête d'optimisation de route envoyé par le MN vers son HA. Il serait ainsi possible à un MN légitime mais malicieux de rediriger les flots vers une victime ou un réseau.

Pour contrer ces difficultés, le protocole Hierarchical Certificate Based Binding Update (HCBU) (Ren *et al.* (2006)) propose une certification en chaîne à trois niveaux où la racine (connue par tous) (niveau 1), les fournisseurs de services (niveau 2) et le domaine du MN (niveau 3) signent le préfixe des sous-réseaux HoA et CoA du MN. La vérification d'une adresse est faite à travers le certificat qui est validé en remontant la hiérarchie jusqu'à trouver l'autorité de confiance. De plus, le MN doit également prouver être le propriétaire légitime du CoA en soumettant une signature du HA du réseau visité à son HA du réseau mère.

Plus sécuritaire que CBU, le HCBU (ou encore sa version optimisée D. Kavitha (2010)) suppose un déploiement global de l'infrastructure de certification à 3 niveaux qui requiert des modifications structurelles globales et une collaboration étroite entre plusieurs consortiums qui régissent Internet. Étant donné la portée relativement limitée de son usage et l'effort considérable à consacrer à un tel déploiement n'est pas justifié. De plus, dans un souci de minimiser les messages de signalisation, HCBU ne vérifie pas l'existence d'une route bidirectionnelle entre le MN et le CN avant son établissement, ce qui peut devenir problématique notamment si les flots traversent un réseau aussi hétérogène que l'Internet.

2.3.3 Algorithmes de génération cryptographique d'adresses

Dans le but d'offrir une authentification décentralisée sans infrastructure, Aura (2003) propose CGA, un algorithme innovateur permettant à un nœud de générer son adresse en la liant à une clé publique et ainsi en prouver sa possession en fournissant une signature avec sa clé privée. L'idée est de remplir les 64 derniers bits d'une adresse IPv6 avec le résultat (hash-1) venant du hachage de la concaténation du modificateur, de l'adresse du sous-réseau, du compte de collision et la clé publique. Puisque 5 bits sont réservés, le nombre d'opérations de hachage sans collision assurant l'usurpation d'une adresse CGA à partir d'une paire de clé publique/privée est de 2^{59}. Afin d'augmenter cette complexité, l'auteur propose l'extension de hachage (hash-2) dans laquelle le modificateur est incrémenté jusqu'à l'obtention de $16 \times SEC$

bits de 0 suite au hachage de la concaténation du modificateur, 9 octets de 0 et la clé publique, où SEC est un entier entre 0 et 7 représentant le niveau de sécurité désiré. La complexité d'usurpation du CGA passe alors de 2^{59} à $2^{59+16\times SEC}$. Cependant, cet effet bénéfique vient au détriment des opérations de hachage nécessaires pour générer un CGA qui passe de 1 à $2^{16\times SEC}$.

Malgré l'extension de hachage proposée, CGA est vulnérable à l'attaque du compromis temps-mémoire dans laquelle différentes valeurs valides du modificateur dans le hash-2 peuvent être pré-calculées pour un SEC et une paire de clé publique/privée donnés. En effet, les éléments invariants dans le calcul de hash-2 permettent de calculer à l'avance des valeurs du modificateur respectant $16 \times SEC$ bits de 0, venant ainsi annuler la complexité ajoutée à l'usurpation du CGA par l'extension de hachage. Ainsi, en supposant aucune collision dans le calcul du hash-1 avec les composants du hash-2 listés dans la table, il faudrait 2^{59} entrées pour ramener la complexité d'usurper un CGA de $2^{59+16\times SEC}$ à 2^{59}.

L'attaque de rejeu (*replay attack*) est également possible puisque la signature prouvant la possession du CGA n'inclut pas l'adresse source du message. Ainsi, un attaquant pourrait emmagasiner les messages envoyés de sa victime et obtenir son modificateur et sa clé publique (deux composantes divulguées publiquement) pour ensuite changer de sous-réseau, générer un nouveau hash-1 à partir modificateur et la clé publique obtenus de sa victime pour refléter la nouvelle adresse du sous-réseau et y retransmettre les messages emmagasiner.

CGA++ (Bos *et al.* (2009)) est une alternative à CGA conçue dans le but de pallier ses vulnérabilités. Pour contrer l'attaque de rejeu, les auteurs proposent d'incorporer dans le hash-1 une empreinte digitale du modificateur, du compte de collision et de l'adresse du sous-réseau empêchant ainsi de retransmettre les messages dans un autre contexte que celui précisé par les composants de la signature. Cette stratégie est efficace, mais requiert une validation cryptographique asymétrique exigeant une monopolisation des ressources et un temps de calcul non-négligeable. Pour contrer l'attaque du compromis temps-mémoire, CGA++ intègre l'adresse du sous-réseau dans hash-2 et hash-1. Étant donné les 64 premiers bits dédiés au sous-réseau dans CGA, il faudrait 2^{64} tables contenant chacune 2^{59} entrées (en supposant aucune collusion dans le calcul du hash-1 avec les entrées de la table) à un attaquant pour annuler la complexité apportée par l'extension de hachage. Cependant, en pratique, l'attaque du compromis temps-mémoire pourrait simplement cibler les sous-réseaux les plus populaires pour ainsi élargir au maximum le bassin des victimes potentielles en limitant le nombre de tables à pré-calculées. De plus, cette pratique est problématique dans un contexte de mobilité où une unité mobile, limité en ressources, doit régénérer son adresse cryptographique à chaque relève.

2.3.4 Les solutions d'optimisation de route basées sur CGA

Arkko *et al.* (2007) ont repris l'adresse du réseau mère générée de manière cryptographique(CGHoA) qui a été introduite dans CAM-DH (M. Roe (2002)) et CGA-OMIPv6 (W. Haddad (2005)) dans le but d'ajouter une authentification faible décentralisée au RR sans exiger de suppositions irréalistes. Pour limiter les abus que des MNs légitimes mais malicieux peuvent causer en usurpant l'adresse IP de sa victime comme son CoA dans l'envoi d'un BU vers le CN, le RFC4866 propose également une autorisation basée sur le crédit.

Cette solution est cependant limitée à plusieurs niveaux. D'abord, l'authentification faible permet à un attaquant ayant accès aux messages initiaux entre le MN et le CN de se faire passer pour l'un ou l'autre. L'attaque de l'homme du milieu devient alors possible en permettant à un attaquant de générer son propre CGA et de communiquer avec le CN pour ensuite retransmettre ces messages vers sa victime. De plus, les auteurs se contente de mitiger les effets de redirection des flots suite à l'envoi d'un BU vers le CN avec une adresse usurpée plutôt que d'éliminer le problème. En effet, le mécanisme d'autorisation basé sur le crédit se fie sur la confiance établie avec le CN pour rediriger les flots vers la nouvelle adresse spécifiée par le MN. Plus cette confiance est élevée, plus les messages seront transférés vers la nouvelle adresse avant que celle-ci ne soit validée. Autrement, si cette confiance est en deçà d'un seuil, le CN doit vérifier l'accessibilité du MN à cette nouvelle adresse avant de rediriger les flots. Or tout système basé sur le crédit ou la réputation laisse une zone grise permettant l'abus des différences entre le seuil de classes. Cette approche est vulnérable à une attaque de déni de service distribuée qui permettrait à un groupe de plusieurs MNs de se synchroniser pour agir correctement pendant un certain temps et ainsi accumuler des crédits, pour ensuite rediriger les flots simultanément vers une même victime. Pour terminer, puisque la solution proposée dans le RFC4866 est basé sur le RR, l'échange d'un nombre très grand de messages dans le réseau d'accès radio va certainement désenchanter les opérateurs qui peuvent voir leur coût augmenter significativement en adaptant le protocole.

CHAPITRE 3

COLLUSION-RESISTANT REPUTATION-BASED INTRUSION DETECTION SYSTEM FOR MANETS

Angelo Rossi and Samuel Pierre
angelo.rossi@polymtl.ca samuel.pierre@polymtl.ca
Mobile Computing and Networking Laboratory (LARIM)
Ecole Polytechnique de Montreal
Montreal, H3T 1J4 Canada

Abstract

Most intrusion detection systems (IDS) for mobile ad hoc networks (MANETs) are based on reputation system which classifies nodes according to their degree of trust. However, existing IDS all share the same major weakness: the failure to detect and react on colluding attacks. The proposed IDS effectively integrates the colluding risk factor into the computation of the path reliability which considers the number and the reputation of nodes that can compare both the source message and the retransmitted one. Also, the extended architecture effectively detects malicious and colluding nodes in order to isolate them and protect the network. The simulations launched in various MANETs containing various proportions of malicious and colluding nodes show that the proposed solution offers a considerable throughput gain compared to current solutions. By effectively selecting the most reliable route and by promptly detecting colluding attacks, the number of lost messages is decreased, and therefore, offering more efficient transmissions.
Keywords: Intrusion detection system (IDS), mobile ad hoc networks, reputation system, colluding attack, network security

3.1 Introduction

Cooperation enforcement models for mobile ad hoc networks (MANETs) are based either on trust management mechanisms Marti *et al.* (2000); Michiardi et Molva (2002); Buchegger et Boudec (2002); Liu *et al.* (2004); Rebahi *et al.* (2005) or virtual money Zhong *et al.* (2003); Xue *et al.* (2003). The latter give nodes an incentive to well behave by receiving an amount of virtual money with which they pay other nodes to forward its messages or

access distant services. On the other hand, the incentives based on trust entice a node to well behave to keep good relations with its neighbors and thus preventing them to drop messages and become isolated from the network. The degree of trust between nodes is measured through a reputation system by which a node sees its reputation increase after it well-behaved or decrease otherwise. Therefore, the threat of a punishment, resulting in a drop of its reputation, pushes nodes to well behave.

Because every node in MANETs functions not only acts as host but also as a router, the critical operation of forwarding packets may easily be interrupted or corrupted for various reasons, either voluntarily or not. Misbehaving nodes are generally categorized into 3 groups: selfish, malicious and colluding. Selfish nodes main concern is to save as much resource as possible by minimizing the amount of data message forwarding while maintaining a minimum cooperation (above the threshold) to remain in the network. The objective of malicious nodes is to disrupt the network by disseminating false information, overloading neighbors or modifying forwarded messages. Finally, a colluding attack Marshall *et al.* (2003) occurs when two or more selfish or malicious nodes collaborate to make an attack without being detected. By observing actions of surrounding nodes, reputation-based solutions can be quite effective against internal active attacks or selfish behaviors. However, to our knowledge very few IDSs consider colluding attacks Ghosh *et al.* (2005, 2004). Works who do either focus on wormholes attacks Mahajan *et al.* (2008); Su et Boppana (2007) or only work for optimized link state routing protocol (OLSR) Suresh P. *et al.* (2010); Sterne *et al.* (2007).

This paper proposes a mechanism to detect generic colluding attacks while also thwarting them by extending the pathrater component of reputation based IDS. The research goal is to design an intrusion detection system against colluding attacks in MANETs. In Section 3.2, a discussion about the strengths and limitations of current IDSs for MANETs is presented. The proposed solution exposed in Section 3.3 starts by presenting the assumptions and follows with the methods and the algorithms proposed. In Section 4, the experimental results are presented. Finally, Section 5 summarizes the contributions and concludes the paper.

3.2 Existent IDSs for MANETs

IDSs are composed of 3 distinct modules: detection, filter and reputation. The detection module monitors the behaviors of the surrounding neighbors and sends the information to the filter module which reveals events worth noticing. Finally, the reputation module establishes a score system which rewards or punishes the nodes according to the received events. The authors of Marti *et al.* (2000) are the pioneers of the MANETs IDSs with the introduction of the Watchdog and Pathrater scheme. These two techniques significantly improve the

throughput in MANETs in the presence of compromised or malfunctioning nodes.

3.2.1 The watchdog component

The Watchdog component is responsible of monitoring the received messages in promiscuous mode with the purpose of making sure that it has forwarded the message without alteration. Assuming the links are bidirectional (i.e. omnidirectional wireless antennas), when intermediary nodes forward the message to its neighbor, it can also verify that the next hop correctly retransmit the message through the use of the passive acknowledgement. If the message remains unaltered within a specified timeout, the next hop well behaved, else it is misbehaving. According to his behaviors, his reputation will be adjusted. A node can classify a neighbor into one of these 3 classes:

- normal: regroups well-behaving nodes;
- suspect: transitory state for closely monitored nodes;
- malicious: temporary banned and isolated nodes.

Because a node changes state when his reputation reaches a predefined threshold, an attacker can easily exploit the gaps between the thresholds to periodically drop messages. If exploited by several nodes, this simple attack can considerably affect the network throughput. To provide more accurate reputation scores many IDSs propagate indirect observations across the network.

CORE Michiardi et Molva (2002), for example, differentiates incoming reputation alerts by grouping them in 3 distinct classes: direct observations, indirect observations (alerts received from distant neighbors) and functional reputation alerts (behavior to accomplish a specific task). The idea is to increase the accuracy by collecting a larger number of alerts, each having a different weight on a node's reputation depending on its type. CONFIDANT Buchegger et Boudec (2002), which operates similarly to CORE, also accepts indirect observations but their weight is proportional to the reputation of the issuer. Also, in order to reduce the number of alerts crossing the mobile ad hoc network, CONFIDANT only considers negative alerts issued on misbehaviors detections. Note however, the propagation of reputation alerts make these IDSs vulnerable to blackmail attacks.

3.2.2 The pathrater component

In order to mitigate the effects of misbehaving nodes, the Pathrater selects the most reliable route available instead of simply choosing the shortest route. The reliability of a path is obtained by calculating the average reputation of the intermediary nodes.

The presented approaches suffer from collusion attacks where two or more adjacent at-

tackers that are being part of a route collaborate to drop or falsify messages (assuming no encryption is employed). Figure 3.1 illustrates the case where nodes B and C collude to modify message M1 without alerting the source. Consequently, the destination D receives a falsified message while source A remains unaware of an intermediary node's misbehavior.

Figure 3.1 Colluding attack

3.3 The proposed collusion-resistant IDS for MANETs

As depicted in Figure 3.2, the proposed IDS classifies nodes into seven classes: FRESH, PRIVILEGED MEMBER, REGULAR MEMBER, LITE MEMBER, INSTABLE, SUSPECT and BANNED. According to their rank, each node is treated differently. Splitting the MEMBER class into 3 distinguishes well-working nodes from very reliable ones in which more important responsibilities can be assigned.

When a node discovers new neighbors through route discoveries or source routing analysis in which no reputation have been previously assigned, they are moved to the FRESH class. Starting with a rating of 0, a fresh node is closely monitored by its neighbors for a period of T_{NEW} and is not permitted to send its own messages. After the preliminary observations, the node migrates to LITE MEMBER if its rating respects the minimal threshold for that class. In other cases, it heads to the SUSPECT class and his rating resets to 0.

Nodes that are part of LITE MEMBER can fully participate in the network by acting as a host, an intermediary node and also a source. The rating of a Lite Member node must be between REPLITE and REPREG while the number of ascendant transitions between the UNSTABLE and LITE MEMBER classes below TRLITE. The most reliable nodes are part of the REGULAR and PRIVILEGED MEMBER classes. They inherit the functionalities of the LITE MEMBER class and ensures the most delicate tasks such as packet rerouting and the participation in local consensuses (see 3.3.2). By creating 2 classes for reliable nodes, it offers more flexibility with the attribution of important responsibilities while reducing intra-class tolerance abuse. To be part of these classes, nodes must respect the minimum reputation threshold and their number of transitions with lower-rating class must not exceed TRREG or TRPRIV.

When a node's rating goes below REPLITE, it heads to the Unstable transient state where the nodes are temporary placed for reexamination. When entering the UNSTABLE class, their rating is reset to 0 and can only act as hosts and intermediary nodes. After

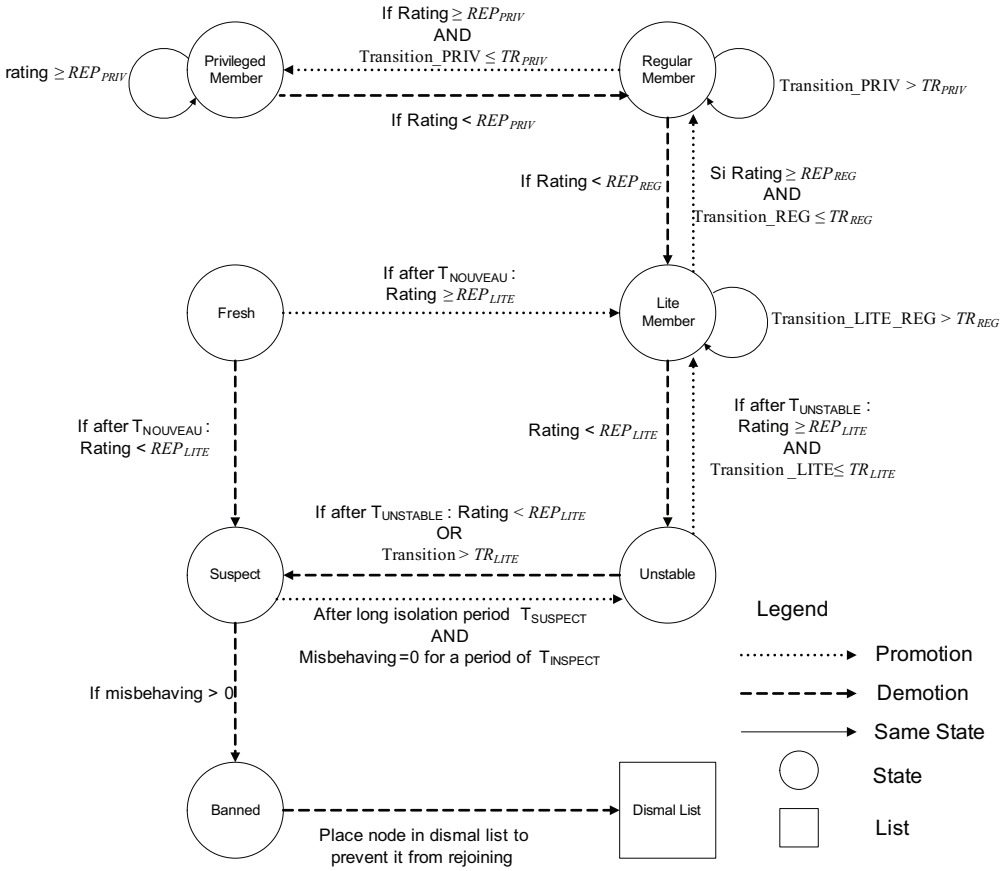


Figure 3.2 State machine diagram depicting the operation of the proposed solution

TUNSTABLE, their condition is reevaluated. If they are unable to upgrade to the LITE MEMBER class (either because their rating is too low or the number of maximum transitions has been exceeded), then they will be temporarily isolated in the SUSPECT class.

Suspect nodes are temporary isolated from the network by not being able to send, retransmit or receive any messages for a period of TSUSPECT. During that time, they will be closely monitored by their neighbors for a period of TINSPECT where any misbehavior will drag the defective node to the permanent dismal list. If no misbehaviors have been noticed, the node will be allowed to reintegrate the network through the UNSTABLE status.

Finally, banned nodes are part of a subjective dismal list maintained by each node individually which forbids them to reenter the network at a later time.

3.3.1 Pathrater and colluding risk factor

All source routing protocol headers provide rich information on the networks topology. The proposed system collects this information and creates a connectivity table which keeps track of the existent links between nodes. Note that if the assumption that all connections are bidirectional stands, the matrix will be symmetric and only half has to be saved in memory. The following model uses the connectivity table in order to wisely choose the most reliable routes by reducing the risk of colluding attacks.

The objective

$$max \sum_{j \in J_p} TL_i(j), i \in I_p, p \in P \tag{3.1}$$

As with other reputation based algorithms, the main objective is to select the path with the highest reputation. However, the computation of the route trust level considers the colluding attack risk factor.

The idea behind the evaluation of the colluding attack risk factor is to determine the number and reputation of available nodes which can detect colluding attacks. Such node must be able to receive messages from an intermediary node $j \in J_p$ and the next intermediary node $j+1 \in J_p$ in the path $p \in P$. These nodes will be called surveillance nodes. The partial trust level of intermediary node j calculated by the source node i who is seeking the safest route to reach the destination is given by :

$$TL'_i(j) = n_j \sum_{k \in K_j} \sum_{z \in (K_k \cap K_{k+1})} \frac{w_{TL_z(k)} TL_z(k)}{|K_k \cap K_{k+1}|} \tag{3.2}$$

The first and most important criterion is the respect of the minimal number of surveillance nodes NminKi between the current and the following intermediary node. Once that number is reached, the average of the gathered trust levels about the surveillance nodes is evaluated.

Table 3.1 Primary factors

Sets	
N	Set of all nodes
P	Set of all nodes forming a path from the source to the destination, $P \subseteq N$
I_p	Set of source nodes for path $p \in P, \lvert I_p \rvert = 1$
J_p	Set of intermediary nodes for path $p \in P, \lvert J_p \rvert = \lvert P \rvert - 2$
K_i	Set of neighbors to $i \in N$ (where $\lvert K_i \rvert$ is the cardinality of K_i representing the number of neighbors of i)
Constants	
TL_iMIN	minimal allowed trust level for $i \in N$
$Nmin_{K_i}$	minimal number of $\lvert K_i \rvert, i \in N$
TL_iMIN_{ALERT}	minimal trust level to consider an alert
$ME_i(j)$	maximum tolerance margin between the trust level received from node $i \in K_j$ for node $j \in N$ and the current trust level
Variables	
$TL_i(j)$	trust level of node $j \in N$ from node $i \in N$ point of view
$w_{TL_i(j)}$	0-1 variable such that $w_{TL_i(j)} = 1$ if and only if $TL_i(j) \geq TL_iMIN, i, j \in N$
n_j	0-1 variable such that $n_j = 1$ if and only if nodes $\lvert K_j \cap K_{j+1} \rvert \geq Nmin_{K_i}, j \in J_p$
w_{ij}	0-1 variable such that $w_{ij} = 1$ if and only if $TL_i(j) \geq TL_iMIN_{ALERT}$ and $\left\lvert \frac{TL_j(x) - TL_i(x)}{TL_i(x)} \right\rvert \leq ME_i(j), i, j, x \in N$

In order to demote surveillance nodes with reputation reports below TLiMIN, trust levels below that threshold are brought down to 0. Note that this strategy requires second hand information to be exchanged among nodes, making this protocol vulnerable to blackmail attacks. A simpler approach would be for the source nodes i to rely only on their own observations to evaluate the surveillance nodes k and thus eliminate trust level alerts propagation across the network. Equation 3.4, similar to the previous, shows the algorithm add-on to the pathrater which takes the risk of colluding attacks into consideration.

$$TL'_i(j) = n_j \sum_{k \in (K_j \cap K_{j+1})} \frac{w_{TL_i(k)} TL_i(k)}{|K_j \cap K_{j+1}|} \tag{3.3}$$

Based on Liu *et al.* (2004); Rebahi *et al.* (2005) in an environment which permits trust level alerts propagation, the complete trust level can be evaluated with:

$$\begin{aligned} TL_i(j) = & \frac{1}{|K_j|} \sum_{k \in K_j} \left(\frac{W_{TL_i(k)} \times TL_k(j) \times CR_i(k) \times (HMAX_i - H_i(k))}{CMAX_i \times HMAX_i} + \right. \\ & \left. n_j \sum_{z \in (K_k \cap K_{k+1})} \frac{w_{TL_z(k)} TL_z(k)}{|K_k \cap K_{k+1}|} \right) \end{aligned} \tag{3.4}$$

By denying trust level alerts forwarding, many simplifications can be applied. First, the factors involving the number of intermediary nodes who forwarded the alerts may be eliminated. Also, because a local consensus (see 3.3.2) is processed every time a neighbor node changes state, it is not necessary to consider the aging on the alerts. Such concept may still be employed in case some nodes misses the direct alerts from their neighbors, but the probability is generally very slim unless they are overcharged or the network is locally congested. However, such concept makes it difficult for a node to have an accurate precision on foreign nodes. The introduction of the factor 1/Hi(j) palliates this uncertainty by minimizing the influence of those reputations on the path selection. Of course, this factor must be chosen wisely to avoid selecting the path only on the reputation of the closest next intermediary node. The pathrater equation 3.6 is therefore simplified in two terms which can be calibrated by introducing the α and β coefficients. Note that $\alpha + \beta = 1$ and $\alpha, \beta > 0$. These coefficients vary in function of the network's security objectives and on available information on the network and remain constant.

$$TL_i(j) = \alpha \frac{TL_i(j)}{H_i(j)} + \beta n_j \sum_{k \in (K_j \cap K_{j+1})} \frac{w_{TL_i(k)} TL_i(k)}{|K_j \cap K_{j+1}|} \tag{3.5}$$

3.3.2 Local Consensus

Reputation based IDSs rely on the quality and the quantity of gathered information to select the most reliable path available. Many IDSs such as CONFIDANT therefore permit foreign nodes to spread their direct observations across the network. Such technique definitely increases the quantity of available information, but does not validate the quality and makes it more vulnerable to blackmail attacks. In order to reduce congestion and increase traffic efficiency, while still maintaining adequate reputation accuracy, nodes initiate the local consensus upon a node migration to another state. By first updating the reputation of the concerned node, all neighbors exchange the reputation of the evaluated node to reach a consensus.

By limiting the propagation of the trust level alerts to their direct neighbors, blackmail attacks will be hard to conduct while still keeping great accuracy locally on nodes reputation. In fact, all gathered reputation alerts are first validated to make sure that the reputation on the concerned node is not too far from the one evaluated with direct observations. Once accepted, it is also pondered with the reputation of source node. However, the lack of information on foreign nodes deeply lowers the accuracy in the path reliability evaluation for the source nodes. Consequently, reliable nodes mandated to select the most reliable routes many will execute many path redirections. The formal model follows.

As a simple prevention against blackmail attacks, only alerts issued from nodes above a specified trust level will be accepted. The variable wij discards trust level alerts from unreliable sources with a trust level below TLiMIN. Note that strategy may not be the most accurate. In fact, honest overloaded nodes will usually see their trust level decrease for not forwarding messages and consequently be rejected from local consensuses. On another hand, overloaded nodes will most likely miss a lot of observations on their neighbors and their assigned trust levels will most likely be outdated, therefore acting similarly to blackmail attacks.

As a second attempt to avoid blackmail attacks, gathered trust level alerts must not differ too much from the current node reputation. This criterion is expressed by the second term. Although this article declares the tolerance margin as a constant decided preliminarily by the network administrator, it would be more adequate to be dynamic according to network factors such as the local traffic.

When a node detects a state transition among one of his neighbors, it locally broadcasts a trust level alert specifying the concerned node and his new state. In order to lower the risk of blackmail attacks, only first order observations are exchanged. When a node receives and accepts a trust level alert, it updates the reputation of the concerned node according to equation 3.6. If this pushes the node to migrate to a different state, it informs his neighbors.

Upon all gathered alerts, the trust level updates are weighted by the reputation of the sources which participated in the local consensus. Because the current node i has a perfect score according to his behaviors, it will have more influence in the update.

$$TL_i(x) = \sum_{j \in K_i} \frac{w_{ij} \times TL_i(j) \times TL_j(x)}{\sum_{j \in J} w_{ij} \times TL_i(j)}, x \in K_i \tag{3.6}$$

3.4 Simulation results and analysis

3.4.1 Simulation design

The experiments have been conducted using Qualnet 4.0 to evaluate the proposed solution by comparing its application layer's throughput with the one of the dynamic source routing (DSR) protocol and the watchdog and pathrater (WDPR) IDS. The chosen primary factors with their respective levels are illustrated in Table 3.2 and the simulation details are showed in Table 3.3.

Table 3.2 Primary factors

Factors		Levels	
Name	Symbol	Name	Description
Number of nodes	N	High	60 nodes
		Average	35 nodes
		Low	15 nodes
Percent of malicious nodes	A	High	35%
		Average	20%
		None	0%
Percent of colluding attackers	C	High	70%
		Average	35%
		None	0%
Mobility	M	High	Low : 2mps High : 20mps
		Average	Low : 0mps High : 10mps
		None	immobile

3.4.2 Results and Analysis

Figure 3.3 plots the average throughput in an environment with 20% of selfish nodes and 35% of colluding attackers. Results show that the network, exposed to the defined environment, should be composed of at least 25 nodes to get a decent throughput. Because

Table 3.3 Simulation details

Static factor	Description	
Simulation Time	3 minutes	
Terrain	2km × 2km	
Number of executions per scenario	30 different seeds	
Application	protocol	CBR
	Number transfers	2
	throughput	4096 bps
Node position	Random	
Node direction	Random waypoint	

the territory surface is a lot bigger than the wireless broadcasting range, enough nodes should be available in order to find a route to the destination. On the other hand, a high density leads to interferences causing a high collision rate, thus explaining the slight throughput fall when the network totals 60 nodes. Most interestingly, the proposed solution outperforms DSR and WDPR especially in a medium size network. In fact, WDPR and DSR are vulnerable to colluding attacks and thus greatly affected when the number of safe colluding-free routes is limited.

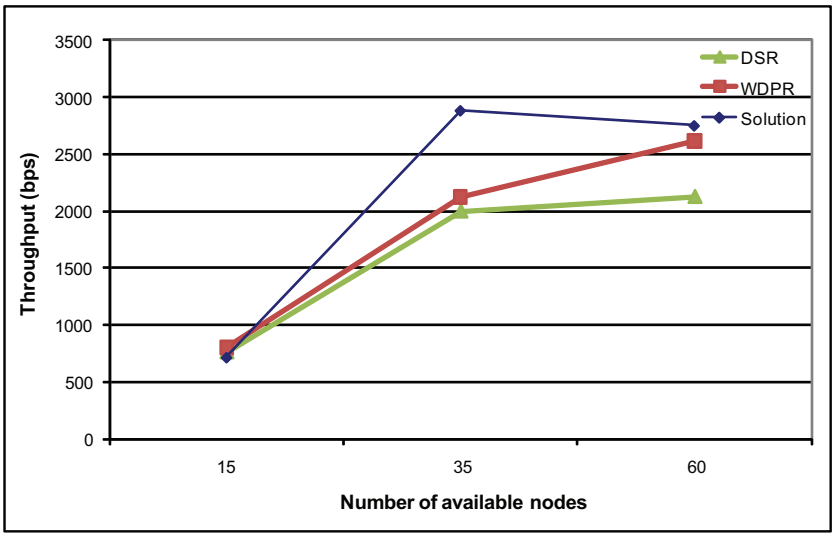


Figure 3.3 Overall throughput as a function of the number of available nodes in the network (20% malicious and 35% colluding)

Figure 3.4 compares the throughput under various percentages of selfish nodes who drop messages at a specified rate. The network has been configured with 60 fixed nodes and no colluding attackers. Results show that the throughput tends to stabilize with the increasing number of selfish nodes for the proposed IDS while decreasing linearly for DSR and WDPR.

This can be explained by the rerouting feature which permits highly reliable surveillance nodes to select an alternative path to reach the destination without having the source to wait for his timeout to expire or for an alert to reach him. However, if most surveillance nodes are selfish, they will not reroute the messages and thus the throughput will rapidly decrease.

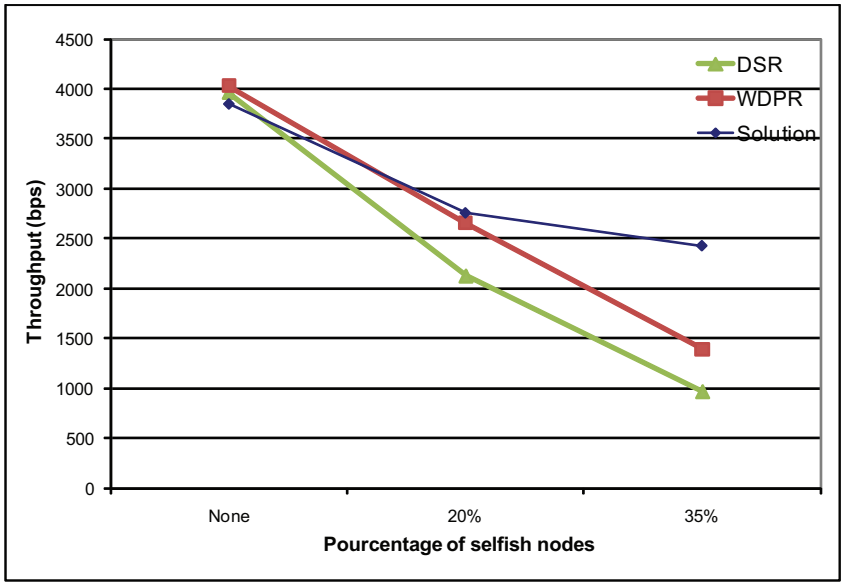


Figure 3.4 Overall throughput as a function of the percentage of selfish nodes in the network (60 fixed nodes and no colluding attackers)

As expected, Figure 3.5 shows that the new proposed security mechanisms developed to counter colluding attacks perform almost independently of the ratio of colluding attackers (but below a given threshold). It also shows how WDPR is vulnerable, reaching the same throughput as the DSR protocol with no security mechanisms in place. One could argue that 70% of colluding attacks is not a realistic scenario. It is important to note that a colluding attack occurs when a colluding attacker precedes a selfish node. Therefore, a presence of 70% of colluding attackers in a 20% selfish acting nodes means that there is a 14% (20% $\times$ 70%) risk of a colluding attack to happen. Also, in order to affect the throughput, the colluding attack must occur in a selected path to reach the destination. If the percentage of colluding attackers is too high, the lack of legit surveillance nodes will corrupt our security mechanisms by always selecting the worse path. On the other hand, if the ratio is too low, there will simply not be any colluding attacks in our simulations.

Ad hoc networks are distinguished from other mobile networks by his dynamic topology. As depicted in Figure 3.6, the delay of convergence for the routing protocol to adapt from sudden route failures and topology changes will inevitably have negative impacts on the network's throughput.

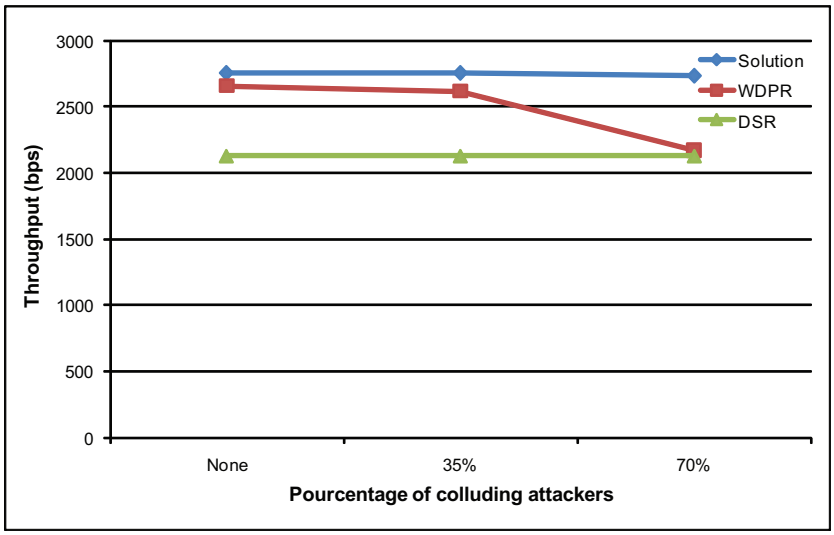


Figure 3.5 Overall throughput as a function of the percentage of colluding attackers in the network (60 fixed nodes and 20% selfish attackers)

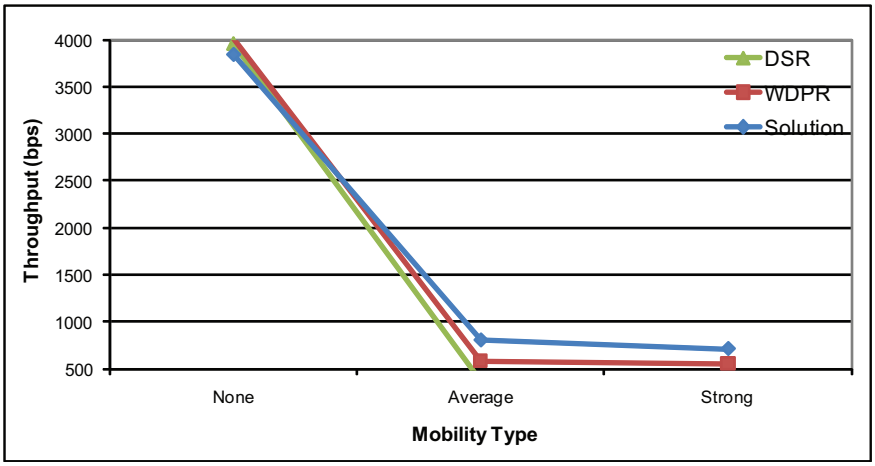


Figure 3.6 Overall throughput as a function of the nodes mobility (60 mobile nodes and no attackers)

Many factors contribute to the important throughput fall. First, the topology, maintained locally in each node, updates via the reception of route error alerts. However, in such an environment, these error messages may not reach the source and therefore it remains unaware of any changes. Because the functionalities of DSR, from which many IDSs are based on, are dependant on accurate topology information, mobility deeply affects our IDS. Second, the continuous arrival of new nodes and the departure of neighbors also influence the throughput in two ways. Because IDS' performance is directly related to the acquired knowledge on active nodes in the network, nodes that briefly cross a neighborhood will not get properly classified by his peers. Because new nodes are classified as FRESH, the network will not be able to take advantage of the legitimate nodes in the participation of local consensus or on the detection of colluding attacks. Figure 3.7 illustrates how efficient are the new security mechanisms to counter colluding attacks in a very hostile environment with a high colluding attack risk (24.5%). First, the local consensus enables a fast detection of misbehaving nodes. Also, the pathrater selects the safest route based not only on nodes' reputation, but also on the current network's topology to exploit surveillance nodes to react on colluding attacks by retransmitting the lost message via an alternative route.

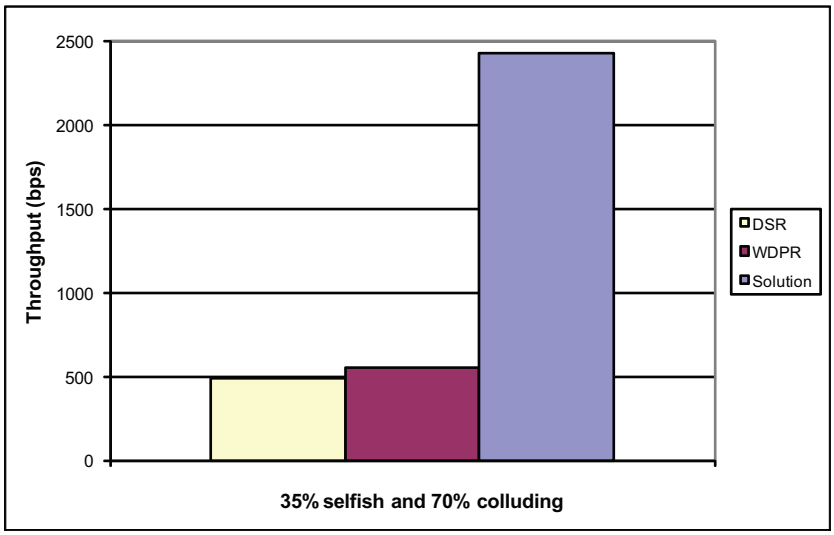


Figure 3.7 Protocols comparison in a hostile environment (60 fixed nodes, 35% selfish and 70% colluding)

3.5 Conclusion

In this paper, we have presented a new pathrater algorithm evaluating the reliability of routes not only by the reputation of the intermediary nodes, but also by the number and

reputation of available surveillance nodes. Their scope is to monitor transmissions between two adjacent intermediary nodes and detect colluding nodes. If they are part of highly reliable classes, they can also reroute the original message through another path. Also, the local consensus technique enables an effective reputation evaluation by exchanging alerts with direct neighbors which greatly reduces communication overhead as compared to the schemes that maintain global reputation.

Simulations results show that colluding attacks do not affect the proposed IDS as much as the original watchdog and pathrater. In fact, the overall network throughput remained constant with the arrival of colluding attackers while decreasing drastically with WDPR and DSR.

On the other hand, the practice of choosing the paths with a higher node density will also increase the risk of transmission collisions due to an environment more conducive to internode interferences. It is safe to conclude that the proposed IDS performs better than others in a realistic mobile ad hoc scenario with lightweight traffic.

An interesting future work would be on implementing a dynamic calibration based on statistical analysis to automatically determine optimal threshold values. The adaptative solution will therefore optimize the parameters for the pathrater such as the rating increment and decrement amounts, the timeout delays and isolation times, and the affected weight on the colluding factor. Many QoS attributes of the networks will be gathered and shared among peers to dynamically set these thresholds to their optimal values and consequently increase throughput regardless of the changing conditions affecting the network.

CHAPITRE 4

AN EXTENSIBLE GAME-THEORETIC FRAMEWORK BASED ON BERTRAND COMPETITION FOR QOS SUPPORT IN MANETS WITH UNCERTAIN ASYMMETRIC COSTS

Angelo Rossi and Samuel Pierre
angelo.rossi@polymtl.ca samuel.pierre@polymtl.ca
Mobile Computing and Networking Laboratory (LARIM)
Ecole Polytechnique de Montreal
Montreal, H3T 1J4 Canada

Abstract

Thwarting selfishness in MANETs requires the integration of reputation-based or nuglets-based (virtual money) incentive mechanisms in the routing protocols to stimulate cooperation between nodes. This paper presents a game-theoretic framework based on Bertrand competition with firms having asymmetric costs and access imperfect information to incite relaying nodes in forwarding messages according to QoS requirements. For a source to send or access QoS-sensitive flows, such as real-time applications, it starts by sending a contract specifying the QoS requirements, its duration and a reservation price. Upon receiving a contract submission, intermediary nodes forming a route between the source and the destination share their current and past collected information on themselves and on surrounding nodes to estimate the probability of breaching the contract and the number of active competitors, and then set a price according to both parameters. the source then selects the cheapest route. The mixed-strategy equilibrium indicates positive industry profits which decline not only with the number of firms having an estimated cost below the reservation price but also with the perception of a greater accuracy on a player's cost that competitors have. Results show that cost uncertainty increases firms' gross margin rate and the prices fluctuation while making the contract honoring much riskier.

Keywords: Bertrand competition, asymmetric costs, uncertainty, imperfect information, mixed strategy, game theory, mobile ad hoc networks, quality of service

4.1 Introduction

Two distinct security objectives in ad hoc networks which must be satisfied in order to assure the network's well behaviour: thwart malice and mitigate selfishness. The former involves using security primitives to ensure integrity, confidentialityCapkun *et al.* (2003); Ng et Seah (2003); Miyao *et al.* (2008) and authenticationHu *et al.* (2005) for features such as naming or addressing of nodes and neighbour discovery protocols. On the other hand, resources scarcity in MANETs leads to selfishness. Such a behaviour disrupts the network functionalities with unfair bandwidth sharing and unwilling packet forwarding. Although these issues cannot be eliminated, they can be mitigated with the use of incentives to stimulate intermediary nodes to forward packets on behalf of other nodes Afergan (2006); Zhou *et al.* (2004); Lu et Pooch (2004); Anderegg et Eidenbenz (2003); MacKenzie et DaSilva (2006); Michiardi et Molva (2003); DaSilva et Srivastava (2004); Xiao *et al.* (2005); Yoo et Agrawal (2006); Srivastava *et al.* (2005).

These incentives can take the form of either reputation systems Buchegger et Boudec (2002); Jaramillo et Srikant (2007); Michiardi et Molva (2002) or virtual monetary transfers Qiu et Marbach (2003); Eidenbenz *et al.* (2008); Xue *et al.* (2003); Zhong *et al.* (2003). Reputation-based solutions collect statistics from surrounding nodes with direct and possibly second-hand (indirect) observations from which they set a reputation (score or rating) for every nodes they discover. The source then picks the route with the highest reputation average. On the other hand, the virtual money incentive works by remunerating (rewarding) the intermediary nodes for their forwarding service with the sender's money which in turn can be used to send their own packets or access a service.

Routing protocols based on pricing schemes have been extensively studied for classic networks in the past Dasilva *et al.* (2000); Lazar *et al.* (1997). More recently, authors in Qiu et Marbach (2003) propose a non-cooperative iterative game where each node selects the optimal bandwidth they can allocate to a source for which it sets a price according to its external demand and remaining bandwidth and energy. However, such game design poses issues for real-time applications that require stringent QoS requirements that only sources are aware of. Also, the lack of communication pushes intermediary nodes to over-allocate bandwidth since the route's maximum bandwidth is limited to the lowest allocation among intermediary nodes. Bandwidth over-allocation not only increases the price for the relay service, but also causes a higher discarded service requests rate because of the lack of unallocated resources. Finally, the authors fail to address the interference issues or consider other important MANETs characteristics such as mobility and radio reliability. On that note, authors in Lu et Pooch (2004) extended the Nash Bargaining Solutions to support

multiple players and found that the equilibrium is reached when compensating equally the intermediate nodes and their neighbours. A more classical cost-efficient routing protocol design is detailed in COMMITEidenbenz *et al.* (2008). Similar to Ad Hoc-VCG Anderegg et Eidenbenz (2003), it covers the route discovery and the packet forwarding phases in which the players optimal strategies is to say their true private values (truthfulness property).

This paper presents a game-theoretic framework based on Bertrand competition with uncertain asymmetric costs aiming at supporting QoS for real-time applications in MANETs. Section 4.2 details Bertrand competition models under various costs and uncertainty assumptions. The QoS framework under Bertrand competition with asymmetric costs and uncertainty follows in section 4.3. The empirical results and the conclusion completes the paper in sections 4.4 and 4.5.

4.2 Background and related work

Bertrand competition models have been studied with various cost assumptions. When two symmetric firms compete by simultaneously setting prices and the market demand possesses a finite choke-off price, the famous "Bertrand paradox" reaches the unique Nash equilibrium when the price equals the marginal cost for zero profit earnings. If no finite choke-off price exists and monopoly revenues are unbounded, authors in Kaplan et Wettstein (2000) show that on top of the pure strategy zero profit Nash equilibrium, there exists a continuum of atomless mixed strategy equilibria in which firms earn positive profits. Others in Dastidar (1995); Hoernig (2002) have showed the existence of pure and mixed strategies for decreasing returns to scale cost functions (strictly convex costs) while Baye et Kovenock (2008) proved that no equilibrium exists for concave costs unless the "winner-takes-all" sharing rule is adopted Vives (1999).

When firms have asymmetric costs and constant marginal costs, Blume (2003) showed that the equilibrium in a duopoly is reached when the low-cost firm sets a price equal to the higher marginal cost while the high-cost firm uniformly randomizes its price on a interval above. Author in Marquez (1997) extended the game for multiple players with asymmetric fixed costs and concluded that only the two most low-cost firms enter the market.

One last extension to the Bertrand model is the cost uncertainty, that is when firms are aware of their own cost but are unsure of the cost type of their rivals. Authors in Spulber (1995); Vives (1999) showed that a Bayesian Nash equilibrium always exists when firms have constant marginal costs drawn from a continuous probability distribution with compact support. Moreover, if the cost function is parameterized with the uncertainty parameter drawn from a continuous probability distribution, firms make positive profits by pricing above

marginal cost. A mixed strategy for discrete cost uncertainty also exists as demonstrated by the author in Routledge (2010).

Competition uncertainty Janssen et Rasmusen (2002) in which a firm may be inactive with a known probability has also been integrated into the Bertrand pricing model. Given a constant activity probability for each firm, the model has a mixed-strategy equilibrium where industry profits are positive and decline with the number of firms. This result is intuitive as less competition typically leads to higher prices until reaching monopoly. First note that the model assumes that the probability of a firm entering the market is kept constant and independent of the price. This assumption is unrealistic as a firm becomes active when the price is above its marginal cost. Second, the model only considers symmetric firms who share the same marginal cost.

Authors in Lizhi Wang et Valenzuela (2007) consider the Bertrand, Cournot and Supply Function Equilibrium oligopoly models for the electricity bidding market with demand uncertainty . Although their results cannot be transposed for MANETs, they are the pioneers in integrating the unit reliability in a Bertrand competition pricing model. The mobility and node reliability for MANETs have been modelled by Cook (January 2009). Note however, that the unit reliability modelling has been oversimplistically reduced to a Weibull distribution, neglecting the individual impact of interferences and radio reliability.

4.3 The QoS framework under Bertrand competition with asymmetric costs and uncertainty

4.3.1 Game description

Assumptions

The proposed framework relies on the following assumptions:

- the existence of a tamper-proof hardware or smartcard on each node to securely manage virtual money and to only transfer money from the sources to the intermediary nodes if the contract is not breached for its entire duration;
- each node behaves rationally to maximize its utility function;
- nodes do not lie when disclosing private information about themselves or past collected information;
- no collusion exists between sources and destinations to lie about the number of competitors and the identity of the intermediary nodes.

Modeling routing as a two-step game

The game is modelled as a Bertrand oligopoly where the sources are the consumers looking for a route at the cheapest possible price. On the other hand, the intermediary nodes forming disjoint routes between the source and the destination are the firms who want to maximize their profits. The proposed routing protocol design extends the dynamic source routing protocol D. Johnson (2007) to allow a two-step game.

First, a source publishes a contract defining its QoS requirements and duration along with the reservation price (RP) in a route request (RREQ) to reach a destination. Upon receiving the RREQ, the intermediary nodes first verify if they can meet the QoS requirements and if so, they append their identity along with their own and others private information to share and then broadcast the modified RREQ until reaching the destination. After waiting a timeout to receive all RREQs on a given contract, it creates a route reply (RREP) for each RREQ and appends not only all the information that resided in the RREQ but also the number of disjoint routes with the identity of the intermediary nodes composing them. Upon sending the RREP back to the source, a consensus on the price to set occurs among the intermediary nodes according to the available information in the RREP.

Second, the source selects the cheapest route among the received responses and pay the winning intermediary nodes by splitting the bid price equally among them (excluding the destination). Note that if non-disjoint routes exist for the same contract, the common intermediary nodes only select the route with the cheapest cost. Otherwise, nodes will create competition and thus lowering their gross margin (GMR) and utility, going against a rational behaviour. If two or more non-disjoint routes have the same lowest cost, one is chosen randomly. Therefore, each RREP received by the source is linked to a disjoint route (firm).

The proposed framework limits the QoS requirements to bandwidth allocation only. It is however meant to be extended to include many other QoS metrics. The information exchanged between nodes is thus limited to:

- radio reliability shape and scale of the Weibull distribution;
- residual energy and remaining nuglets (money) about themselves and about past and current neighbours;
- average interference and its related number of observations and variance;
- mobility volatility information and current and past neighbours.

Formal game characteristics

More formally, the type of the game has the following characteristics:

- Bertrand oligopoly : the competition is purely based on pricing where the sources

(consumers) pay a coalition of intermediary nodes (oligopolists) forming a path to the destination for relaying the data according to QoS requirements defined by the sources;
- coalition-cooperative : the coalition of nodes within a player share private information among them through RREQs to set a price for the contract (there is no cooperation between players);
- asymmetric : players have different payoffs and costs due to variable private information among nodes.
- simultaneous : players set a price simultaneously for a given contract relying only on past private information of a subset of nodes;
- imperfect information : players rely on past collected data on various nodes and therefore their beliefs on the accuracy of the private information leads to a Bayesian mixed strategy Nash equilibrium.

4.3.2 Notations

The sets are:

N: set of all the nodes in the ad hoc network;
I: set of all subsets of nodes ∈ N forming disjoint routes (firms) between the source and destination, exclusive ∈ N;
K: set of all knowledge variables;
C: set of all contracts;
Q: set of all QoS requirements.

The variables are:

$c^q_{\Delta t}$ contract defined by a duration $\Delta t \in t$ and q ∈ Q;
p_i price set by player i ∈ I;
k^i_n knowledge (information) acquired on node n ∈ N by player i ∈ I;
k^i_j knowledge acquired on player j by i (i,j∈ I);
B_n Amount of nuglets in bank for node $n \in N$;
E_n Residual energy for node $n \in N$.

The functions are:

$C_n(c^q_{\Delta t}) \equiv C_n$	The expected cost infringed to node n $\in$ N for respecting contract $c^q_{\Delta t} \in C$ for its entire duration ΔT;
$C_i(c^q_{\Delta t}) \equiv C_i$	The expected cost infringed to player i $\in$ I computed as $\|i\| \times \max(C_n) \forall n \in i$;
$P[BC(c^q_{\Delta t})_n]$	CDF representing the probability for node $\equiv P[BC_n]$ n $\in$ N to fail in respecting the contract $c^q_{\Delta t} \in C$ for its entire duration ΔT ($0 \leq P[BC(c^q_{\Delta t})_n] \leq 1$);
$P[BC(c^q_{\Delta t})_i]$	Cumulative distribution function on the $\equiv P[BC_i]$ breach of contract $c^q_{\Delta t} \in$ C probability for player i $\in$ I;
$E_n(c^q_{\Delta t})$	Expected required energy for node n $\in$ N to honor contract $c^q_{\Delta t}$ for its entire duration Δt;
$\alpha^i_n(p, c^q_{\Delta t}, k^i_n)$	Cumulative distribution function (CDF)$\equiv \alpha^i_n(p)$over p on the probability for node n $\in$ N to participate in contract $c^q_{\Delta t} \in$ C according to player i based and its knowledge $k^i_n \in K, i \in I$ (equivalent to $Pr[C_n \leq p]$) ($0 \leq \alpha^i_n(p) \leq 1$);
$\alpha^i_j(p, c^q_{\Delta t}, k^i_j)$	Cumulative distribution function (CDF) $\equiv \alpha^i_j(p)$ over p on the probability for player j $\in$ I to participate in contract $c^q_{\Delta t} \in$ C according to player i based and its knowledge $k^i_j \in K, i \in I$ (equivalent to $Pr[C_j \leq p]$) ($0 \leq \alpha^i_j(p) \leq 1$);
$F_i(p, \alpha_i(p))$	Mixing CDF on the probability for player $\equiv F_i(p)$ i $\in$ I to offer a better or equal price than p (equivalent to $Pr[p_i \leq p]$) ($0 \leq F_i(p) \leq 1$);
$\pi_i(p, c^q_{\Delta t}, k^i_j)$	Expected profit for player i $\in$ I by setting a price p for the contract $c^q_{\Delta t} \in$ C and knowledge $k^i_j \in K_i, \forall j \in I \setminus \{i\}$;
$U_n(p, B_n, E_n, c^q_{\Delta t})$	Expected utility for node n $\in$ N given its private information, price p and contract $c^q_{\Delta t} \in$ C.

4.3.3 Node preferences

Property 1: The utility function monotonously increases with the selected price

$$\frac{\partial U_n(p, B_n, E_n, c^q_{\Delta t})}{\partial p} > 0 \tag{4.1}$$

Property 2: The utility function monotonously increases with the accumulated nuglets and the remaining energy

$$\frac{\partial U_n(p, B_n, E_n, c^q_{\Delta t})}{\partial B_n} > 0 \tag{4.2}$$

and

$$\frac{\partial U_n\left(p, B_n, E_n, c^q_{\Delta t}\right)}{\partial E_n} > 0 \tag{4.3}$$

Property 3: The user's marginal utility decreases with the remaining nuglets

$$\lim_{B_n \to \infty} \frac{\partial U_n\left(p, B_n, E_n, c^q_{\Delta t}\right)}{\partial B_n} = 0 \tag{4.4}$$

Property 4: The utility function monotonously decreases with the required energy to satisfy the request

$$\frac{\partial U_n\left(p, B_n, E_n, c^q_{\Delta t}\right)}{\partial E_n\left(c^q_{\Delta t}\right)} < 0 \tag{4.5}$$

Property 1 is justified by the fact that a player always seeks to receive the maximum money for their relay services. Simultaneously, a player also wants to accumulate as much nuglets as possible, increasing their buying power to access more available services or sending their own packets. Property 3 reflects the unsatisfactory nature of humans in which the more one possesses, the more he wants. Finally, energy being a scarce non-renewable resource, property 4 considers the energy consumption.

4.3.4 Node reliability

The node reliability is an important aspect to the proposed model as it enables players not only to compute their own failure probability, but also their competitors'. Let $BC_n(c^q_{\Delta t})$ ($BC_i(c^q_{\Delta t})$) be the event where node n ∈ N (player i ∈ I) breaches the contract $c^q_{\Delta t} \in C$ by not respecting its requirements, the breach of contract probability by only considering the radio reliability modelled using the Weibull distribution with Θ_n and $\beta_n, n \in N$ representing its scale and shape:

$$\begin{aligned} P[BC(c^q_{\Delta t})_n] \equiv P[BC_n] &= 1 - \exp^{-\left(\frac{\Delta T}{\Theta_n}\right)^{\beta_n}} \\ P[BC(c^q_{\Delta t})_i] \equiv P[BC_i] &= 1 - \prod_{n \in i} \exp^{-\left(\frac{\Delta T}{\Theta_n}\right)^{\beta_n}} \ \forall i \in I \end{aligned} \tag{4.6}$$

Interference

The definition of radio spectrum interference in the proposed model is reduced to the bandwidth consumption for which the node is not paid for. In fact, when a node transmit a message, it induces an interference range for which nodes within it are negatively impacted with a lower available spectrum. As with the relative mobility, the nodes' interference is a metric estimated using past collected information but because they are not weighted according to the time they occurred, information older than a threshold must be erased. Let C^n_r be

the random variable representing the remaining capacity for node n ∈ N after accepting the submitted contracts, $\overline{C^n_{int}}$ and $\widehat{\sigma}_{C^n_{int}}$ be the average and variance of the capacity used by interference given X observations, focusing only on interference, $P[BC_n]$ follows a predictive Student's t-distribution CDF:

$$\begin{aligned}
&\frac{C^n_r - \overline{C^n_{int}}}{\widehat{\sigma}_{C^n_{int}}\sqrt{1+\frac{1}{X}}} \sim T^{X-1}_n \\
&P\left[C^n_{int} \leq C^n_r\right] \sim Student_{CDF}\left(T^{X-1}_n, X-1, \widehat{\sigma}_{C^n_{int}}\right) \\
&P[BC_n] = 1 - P\left[C^n_{int} \leq C^n_r\right], \forall n \in i \\
&P[BC_i] = 1 - \prod_{n\in i} Student_{CDF}\left(T^{X-1}_n, X-1, \widehat{\sigma}_{C^n_{int}}\right) \forall i \in I
\end{aligned} \tag{4.7}$$

Mobility

Relative mobility between nodes is measured using the average connectivity volatility to estimate the stability of the link. It is computed by counting every arrival or departure of a neighbour and averaged through time. The memoryless property of the exponential function makes it a good candidate to model mobility. Let T be the random time variable, $\Delta t_{n,n+1}$ the duration of the last unbroken link between node n and the next forwarding node n+1 (n ∈ N) and $\omega_{n,n+1}$ their volatility, by considering only mobility, the failure probability becomes:

$$\begin{aligned}
P[BC_n] &= 1 - P\left[T \geq \Delta t_{n,n+1} + \Delta T \mid T \geq \Delta t_{n,n+1}\right] \\
&= 1 - P\left[T \geq \Delta T\right] \\
&= 1 - \exp^{-\omega_{n,n+1}\Delta T}, \forall n \in i \\
P[BC_i] &= 1 - \prod_{n\in i} \exp^{-\omega_{n,n+1}\Delta T}, \forall i \in I
\end{aligned} \tag{4.8}$$

Probability of failure

Assuming constant shape and scale parameters in the Weibull distribution for all nodes ($\beta_n = \beta, \theta_n = \theta\ \forall n \in N$) and combining the radio reliability, the interference and the mobility results in:

$$\begin{aligned}
P[BC_i] =& 1 - \Bigg(\exp^{-\left(\frac{\Delta T}{\Theta}\right)^{\beta\times n}} \times \\
&\prod_{n\in i} Student_{CDF}\left(T^{X-1}_n, X-1, \widehat{\sigma}_{C^n_{int}}\right) \times \\
&\prod_{n\in i} \exp^{-\omega_{n,n+1}\Delta T}\Bigg) \forall i \in I
\end{aligned} \tag{4.9}$$

4.3.5 Utility and cost functions

For nodes to evaluate the cost of accepting a contract in monetary terms, the utility function is used to compute the minimum price for the node to remain neutral in accepting or not the contract. An example of the utility function in 4.10 that respects the preferences stated in 4.3.3 can be as follows:

$$\begin{aligned} U_n\left(p, B_n, E_n, c^q_{\Delta t}\right) =& \ln\left(B_n + \left(1 - P[BC_i]\right)p\right) + \\ & \ln\left(E_n - \left(1 - P[BC_i]\right)E_n\left(c^q_{\Delta t}\right)\right) \\ & \forall n \in i, i \in I \end{aligned} \tag{4.10}$$

The cost is therefore given by :

$$\begin{aligned} U_n\left(0, B_n, E_n, 0\right) =& U_n\left(C_n, B_n, E_n, c^q_{\Delta t}\right) \forall n \in N \\ \ln\left(B_n\right) + \ln\left(E_n\right) =& \ln\left(B_n + \left(1 - P[BC_i]\right)C_n\right) + \\ & \ln\left(E_n - \left(1 - P[BC_i]\right)E_n\left(c^q_{\Delta t}\right)\right) \\ C_n =& \frac{B_n}{\left(1 - P[BC_i]\right)} \times \\ & \left(\frac{E_n}{E_n - \left(1 - P[BC_i]\right)E_n\left(c^q_{\Delta t}\right)} - 1\right) \\ & \forall n \in i, i \in I \end{aligned} \tag{4.11}$$

It is easy to show that the cost function is strictly convex:

$$\begin{aligned} \frac{\mathrm{d}^2 C_n}{\left(\mathrm{d}E_n\left(c^q_{\Delta t}\right)\right)^2} =& \frac{2B_n E_n}{\left(E_n - E_n\left(c^q_{\Delta t}\right)\right)^3\left(1 - P[BC_i]\right)} > 0 \\ & \forall n \in i, i \in I \end{aligned} \tag{4.12}$$

4.3.6 Bertrand pricing model

Before setting a price, firms are uncertain about the number of other firms that will compete for a given contract. A firm i $\in$ I enters the market only if the price p is equal or above its cost C_i. However, the knowledge a firm has on a potential competitor may not be accurate, thus causing an uncertainty around the rival's cost and market participation. $\alpha^i_j(p, c^q_{\Delta t}, k^i_j) \equiv \alpha^i_j(p)$ represents the CDF over price p of a firm j $\in$ I to be active for a given contract $c^q_{\Delta t} \in$ C I from the standpoint of firm i $\in$ I with the knowledge k^i_j it has on j. If firm j's cost is certain and the price p is below it, than $\alpha^i_j(p) = 0$. On the other hand, if $p > C_j$ than firm i is sure that firm j is active $\alpha^i_j(p) = 1$.

In the context of knowledge uncertainty, a firm estimates the competitors participation

using past collected information on neighbours and shared among the firm's coalition of intermediary nodes. The price interval in which $0 < \alpha_j^i(p) < 1$ represents the uncertainty around the competitor's cost. The uncertainty is therefore reflected by the slope of $\alpha_j^i(p)$.

Similarly to the interference estimation, the cost uncertainty is modelled using the predictive Student's t-distribution where $\overline{C_n}$ and $\widehat{\sigma}_{C_n}$ are the sampling average and the variance cost of node n ∈ i given X_n shared observations among intermediary nodes of firm i ∈ I:

$$\begin{aligned} &\frac{p-\overline{C_n}}{\widehat{\sigma}_{C_n}\sqrt{1+\frac{1}{X_n}}} \sim T^{X_n-1} \\ &\alpha_n^i(p, c_{\Delta t}^q, k_n^i) \sim Student_{CDF}\left(T^{X_n-1}, X_n-1, \widehat{\sigma}_{C_n}\right) \end{aligned} \tag{4.13}$$

$\alpha_n^i(p)$ is the confidence level that the real cost of node n ∈ N for accepting the contract falls in the interval $\left]-\infty, \frac{p-\overline{C_n}}{\widehat{\sigma}_{C_n}\sqrt{1+\frac{1}{X}}}\right]$ according to the knowledge owned by player i ∈ I. Consequently, the higher the price set by a firm i, the higher the probability that the cost of node n ∈ j from a competitor j ∈ I is covered and it participates in the market.

The game however is played with firms where each player is a coalition of possibly many nodes and thus a subset of N ($i \subseteq N \forall i \in I$). Let $|i|$ be the number of intermediary nodes (size of subset i ⊆ N) in a disjoint route (firm i), the cost of a firm (route) is:

$$C_i = |i| \times \max\left(C_n\right) \forall n \in i, i \in I \tag{4.14}$$

Extending the route cost definition, $\alpha_j^i(p)$ is therefore characterized by the node $n_{max} \in j$ in competitor j ∈ $I \setminus \{i\}$ which is believed by player i ∈ I to have the highest cost average $\overline{C_{n_{max}}}$ according to its $X_{n_{max}}$ shared observations with a probability close to 1:

$$\begin{aligned} &n_{max} = \arg\max_{n\in j}\left(\arg\left(\alpha_n^i(p, c_{\Delta t}^q, k_n^i) \approx 1\right)\right) \\ &\overline{C_j} = |j| \times \overline{C_{n_{max}}} \\ &\frac{p-\overline{C_j}}{\widehat{\sigma}_{C_{n_{max}}}\sqrt{1+\frac{1}{X}}} \sim T^{X_{n_{max}}-1} \\ &\alpha_j^i(p, c_{\Delta t}^q, k_j^i) \sim Student_{CDF}\left(T^{X_{n_{max}}-1}, X_{n_{max}}-1, \widehat{\sigma}_{C_j}\right) \\ &\frac{\mathrm{d}\alpha_j^i(p, c_{\Delta t}^q, k_j^i)}{\mathrm{d}p} \sim Student_{PDF}\left(T^{X_{n_{max}}-1}, X_{n_{max}}-1, \widehat{\sigma}_{C_j}\right) \end{aligned} \tag{4.15}$$

Let Win_i be the event that firm i wins the contract given N potential competitors, firm i's expected payoff from accepting contract by charging p_i given the knowledge on competing

firms is:

$$
\begin{aligned}
P[Win_i] =& \sum_{\substack{k=1 \\ k \neq i}}^{N} \left[\prod_{j=k}^{N} \left(1 - \alpha_j^i(p, c_{\Delta t}^q, k_j^i)\right) \times \right. \\
& \left. \prod_{l=1}^{k-1} \alpha_l^i(p, c_{\Delta t}^q, k_j^i) \left(1 - F_l(p)\right) \right] \\
\pi_i(p, c_{\Delta t}^q, k_j^i) =& P[Win_i] \times [(1 - P[BC_i]))) \times p - C_i]
\end{aligned}
\tag{4.16}
$$

The strictly decreasing returns to scale or the strictly convexity of the cost function guarantees the existence of a mixed strategy equilibrium Hoernig (2002). In equilibrium, firm i must be indifferent between all pure strategies and therefore the derivative of equation in respect to p is zero ($\frac{d\pi_i}{dp} = 0 \forall i \in I$).

4.3.7 Example with 2 competing firms

Figure 4.1 illustrates the state of a MANET at a given time from which steps describing how the game is played are detailed. As depicted by Table 4.1, there are 3 available routes, but only 2 of them are disjoint and can be firms. Because the route cost is determined only by the number of intermediary nodes and the maximum node cost, nodes 1&2 choose node 3 which has a cost of 100 as opposed to 120 for node 4. If both nodes 3&4 had cost lower than nodes 1 and 2, then both nodes could have been chosen equally as it would not have impacted the route cost. Table 4.2 shows the computed costs for each node according to the shared knowledge. From firm 2 standpoint, route 1 is composed of 3 nodes where the node with the highest cost is node 3 with 100 and precision of 10 observations and variance of 1. Therefore, route 1 estimated cost is determined by the triplet $(3 \times 100, 10, 1)$ and consequently $a_2^1 = Student_{cdf}(T^9, 10, 1)$ where $\frac{p-100}{1\sqrt{1+\frac{1}{10}}} \sim T^9$. For simplicity, $P[BC_1] = P[BC_2] = 0\%$.

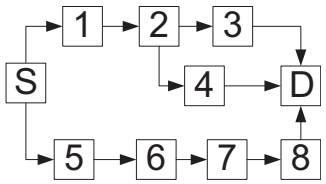


Figure 4.1 Example of a 2-player routing game

For the given network, the proposed framework integrated in the DSR routing protocol follows these steps :

1. A source S sends a RREQ to reach destination D (S,D $\in$ N) at 0.1Mbps for 3 minutes with a reservation price (RP) of 400.

Table 4.1 Available routes (firms) and true cost

Route	Firm	C_n	C_i
1-2-3	1	80-90-100	300
1-2-4	-	80-90-120	360
5-6-7-8	2	60-50-40-80	320

Table 4.2 Resulting $(\overline{C_n}, X, Var(C_n))$ triplet per node estimated from the shared knowledge

Node \ Firm	1	2
1	-	(82,5,2.5)
2	-	(91,7,1.5)
3	-	(100,10,1)
5	(57,2,1.75)	-
6	(50,10,0.5)	-
7	(42,4,1)	-
8	(81,2,4)	-
estimated C_1	-	(300,10,1)
estimated C_2	(324,2,4)	-

2. Intermediary nodes verify that the requested bandwidth is available. If so, they append their identity and private information (radio shape and scale parameters, average interference and variance, mobility volatility, residual bandwidth and energy) on themselves (precise) and other nodes (estimated with average, variance and number of observations) to the RREQ and forward it until reaching the destination D. In this case, the destination receives 3 RREQs.
3. After waiting for the timeout expiration, the destination sends a RREP per received RREQ appending the competing routes on top of all information present in the corresponding RREP.
4. Intermediary nodes compute the cost of respecting the contract for their own route (known) (equation 4.14) and competitors (uncertain) (equation 4.15). Table 4.1 shows the true cost of both firms while Table 4.2 show the estimated cost per node (and firms - last 2 lines) according to firms.
5. Firms now set the price according to the expected profit functions:

$$\begin{aligned}\pi_1(p) &= \left[(1-\alpha_2^1(p)) + \alpha_2^1(p) \times (1-F_2(p))\right] \times \\ &\quad [(1-P[BC_1]) \times p_1 - C_1] \\ \pi_2(p) &= \left[(1-\alpha_1^2(p)) + \alpha_1^2(p) \times (1-F_1(p))\right] \times \\ &\quad [(1-P[BC_2]) \times p_2 - C_2]\end{aligned} \tag{4.17}$$

The equilibrium to this game is :

$$F_1(p) \begin{cases} 0 & \text{for } p \leq p_{min}^1 \\ 1 - \left(\frac{1-a_1^2(p)}{a_1^2(p)}\right)\left[\frac{(1-P[BC_2])[RP-p]}{(1-P[BC_2])p-C_2}\right] & p_{min}^1 \leq p \leq p_{max}^1 \\ 1 & \text{for } p \geq p_{max}^1 \end{cases}$$

$$F_2(p) \begin{cases} 0 & \text{for } p \leq p_{min}^2 \\ 1 - \left(\frac{1-a_2^1(p)}{a_2^1(p)}\right)\left[\frac{(1-P[BC_1])[RP-p]}{(1-P[BC_1])p-C_1}\right] & p_{min}^2 \leq p \leq p_{max}^2 \\ 1 & \text{for } p \geq p_{max}^2 \end{cases} \tag{4.18}$$

$$\text{where } p_{min}^i = RP - \frac{a_i^j(p_{min}^i)\left[(1-P[BC_j])RP - C_j\right]}{(1-P[BC_j])}$$

$$\text{and } p_{max}^i = \min\{RP, \arg(a_i^j(p) \approx 1)\}, i \in I, j \in I \setminus \{i\}$$

Many differences are visible between $F_1(p)$ and $F_2(p)$ illustrated in Figures 4.2 and 4.3. First, the minimum price defined by p_{min}^i for players 1& 2 are respectively 320 and 327. Also, the accentuated slope in F_1 shows that player 1 believes that its rival has an accurate view on

its cost while the player's 2 belief that the risk for player 1 to get its cost wrong is reflected in the slope of F_2. This leads player 2 to pick a price in the interval [327, 400] ($F_2(400) \approx 1$) while player 1 plays the pure strategy $p_2 = C_2 = 320$. The average price of player 2 being 332, the average gross margin rate (GMR) for players 1 & 2 is respectively of 6.25% and 3.6%. The fact that player 2 is aware that firm 1 cost is lower than is explains why it is forced to price closer to its cost.

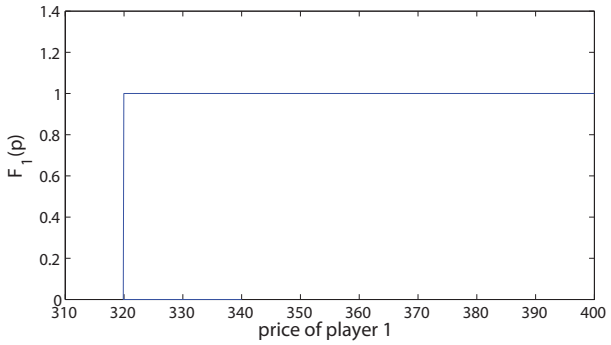


Figure 4.2 Mixing CDF for player 1

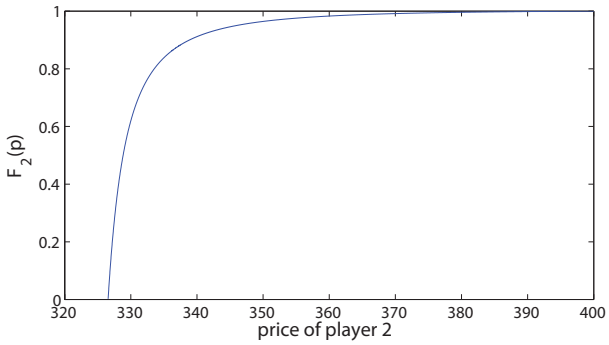


Figure 4.3 Mixing CDF for player 2

The perception on the accuracy of a player's cost by its competitors directly impacts p^i_{min} and $p^i_{max}, i \in I$. The higher its accuracy, the lower the player i will set its price as p^i_{min} and p^i_{max} decreases.

4.4 Analysis and Empirical results

4.4.1 Simulation and empirical results

The simulations have been run on MATLAB r2009b to numerically solve the system of partial differential equations obtained by deriving the profit function in respect to the price variable simultaneously for all players. However, the lack of a MANET simulation block (in Simulink) forced some simplifications in the implementation:

- Signal-related radio characteristics such as fading are ignored;
- Interferences are caused only by neighbours in the transmission range and therefore no hidden node issues are considered;
- All links are bidirectional;
- Mobility does not follow known mobility patterns such as the random way point but is limited in shuffling the global connectivity table every minute.

Table 4.3 shows the details about the simulations. Note that the mobility is expressed in % which indicates the probability of a node to change status with another node. For example,

Table 4.3 Simulation details

Parameter		Value
Simulation Time		until end of contracts
Number of runs per scenario		10
Contracts	number of contracts	5
	Start time	random
	Bandwidth (Kbits)	100, 200, 300, 350 , 400
	Duration (minutes)	3, 2, 5, 4, 3
	Reservation Price	300, 300, 800, 600, 650
Nodes	initial position	random
	radio capacity	1 Mbit
	mobility pattern	random shuffle
	initial nuglets	random in [100 5000]
	initial energy	random in [1000 5000]
	Weibull shape (β)	2
	Weibull scale (θ)	100
Energy consumption (per Kbits*minute)		3
Timeout expiration of information		10 min

for a 5% mobility and a contract of 5 minute duration, if node 1 is currently connected with node 2, there is a $(0.95)^5) = 77\%$ probability that node 1 remains connected with 2 for the entire duration of the contract.

A node's knowledge is characterized by the amount of private information collected on surrounding nodes that have not expired. It is limited by the size of the history buffer and by mobility which enables nodes to have information on more nodes in the network. However, these two factors push prices in opposite directions. On one hand a lower uncertainty leads to lower p^i_{min} and p^i_{max} prices, but on the other, a higher mobility leads to a higher $P[BC_i]$ and thus a higher cost and price. For this reason, the impact of knowledge on the gross margin rate is better reflected through a variable buffer size with a constant mobility rate as shown in Figure 4.4.

Also note that the buffer size also impacts the failure probability which relies on past information on bandwidth to estimate interference. The difficulty in accurately estimating future bandwidth interferences by the lack of knowledge inherently increases the rate of breach of contracts computed as $\frac{\text{number of breached contracts}}{\text{number of accepted contracts}}$. Figure 4.5 shows how a greater knowledge helps players make a better choice to avoid service disruptions.

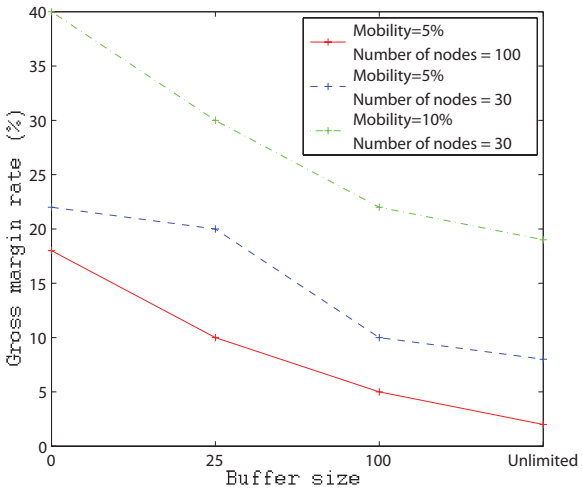


Figure 4.4 Impact of knowledge and number of competitors on the gross margin rate

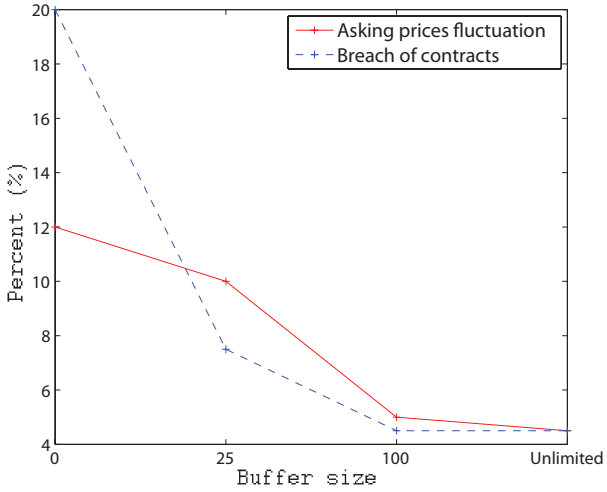


Figure 4.5 Impact of knowledge on price fluctuation and breach of contract (Mobility = 5%, Number of nodes = 30)

4.4.2 Discussion on factors that impact prices

The contract specifications (RP, bandwidth, duration) all have an impact on prices. First, the requested bandwidth determines the energy required to support the contract for its whole duration which inherently impacts the cost of production. Note that the model can be extended to include as many multiple QoS parameters as wanted. However, for specifications that do not impact energy such as time-related QoS parameters like end-to-end delay or jitter for example, a more extensible failure probability ($P[BC_i]$) model is required. In either cases, it impacts the cost and thus adding upward pressure on prices. The influence of the reservation price, the last element of a contract, is easily visible in equation 4.18 as the more it rises, the more p^i_{min} and p^i_{max} increase.

The most interesting factor on the price is the uncertainty on players' competitiveness which affects the slope of the activity probability. A player's best response depends on his belief (or perception) of the amount of knowledge its competitors have about its cost. The more a player believes that competitors have an accurate view on its cost, the more its prices (p^i_{min} and p^i_{max}) fall. Inversely, if a player thinks that others have little or no information about its cost, then it will tend to put higher prices. A secondary effect on the uncertainty is the variation in prices which tends to be higher when uncertainty grows (Figure 4.5).

Finally, a larger number of nodes increases the probability for a node of having a low-cost competitor, thus forcing it to reduce its price to remain competitive. Yet, a node must be aware of such competition to lower its gross margin rate. Figure 4.4 confirms that even with

a network with more than 3 times more nodes in the network, if nodes have little or no knowledge, prices remain high. A steep drop in GMR then occurs when nodes become aware of the increasing competition.

4.4.3 Discussion on the assumptions

One of the core concepts of this framework is uncertainty which is tightly linked with knowledge. Therefore, rational nodes may tend to maliciously falsify or limit the information they disclose. The first debatable assumption is the absence of collusion between sources and destinations. Among the information they use to set prices, firms rely on the destination to disclose the number of competitors along with the identity of the intermediary nodes composing the available routes. As the consumer best interest is to pay as little as possible, sources could collude with destinations to publish inexistent disjoint routes and then publish a false low cost on the involved intermediary nodes (low-cost competitors). The use of cryptographic solutions can partially tackle the issue by forcing nodes to sign their submitted private values (information). However, along with the key management issues in a decentralized environment, the destination could still create inexistent nodes to falsely give the impression of increasing competition. This remains an open issue.

Second, in the proposed framework, a firm is a coalition of nodes where they share private information about themselves and other nodes they communicated with previously. In the case of a monopoly, the price fixed by a firm will always be the reservation price stated in the contract if such price covers the cost. However, when competition reigns, a rational node could tend to lie if by doing so, the contract is still accepted by the source and its GMR increases. To counter this, the framework is designed to compute the cost according to the highest cost among the intermediary nodes of a route and then split the price billed the source equally among intermediary nodes of the winning route. Therefore, if a node discloses a higher cost to force the coalition to set a price higher than what competitors believe its route cost is, chances are that the source selects a cheaper price from a competitor. The lie thus transforms a potential gain into nothing, which goes against a rational behaviour. The same logic holds when lying about competitors' costs.

4.5 Conclusion

This paper presents a framework based on Bertrand competition with uncertain participation for firms with different cost. The model extends existing works by supporting multiple asymmetrical firms and by generalizing the activity probability based on the price. It has been constructed in regards of a routing protocol supporting QoS parameters for MANETs

with the scope of analyzing the expected performance of the network in terms of prices and data flow disruptions when nodes behave rationally. The results show that knowledge on other participants' cost improves the network's welfare by reducing the breaches of contracts while enticing nodes to price closer to their cost.

CHAPITRE 5

AN EFFICIENT AND SECURE SELF-HEALING SCHEME FOR LKH

Angelo Rossi and Samuel Pierre
angelo.rossi@polymtl.ca samuel.pierre@polymtl.ca
Mobile Computing and Networking Laboratory (LARIM)
Ecole Polytechnique de Montreal
Montreal, H3T 1J4 Canada

Suresh Krishnan
suresh.krishnan@ericsson.com
Ericsson Research
Town of Mount Royal, QC H4P 2N2, Canada

Abstract

With the growing interest in converging fixed and mobile networks (FMC), mobile applications will require more and more resources from both the network and the mobile unit. In such context, multicasting is essential because it lowers bandwidth consumption by simultaneously reaching a group of multiple recipients. Securing multicast flows has been extensively studied subject in the past decade, but none of the existing solutions were meant to handle the constraints imposed by mobile scenarios, in particular the high packet-loss rate. The need for a low overhead self-healing rekeying distribution that is scalable, reliable and suitable for mobile environments has never been more urgent than with the arrival of FMC in 4G networks. This paper presents two self-healing recovery schemes based on the dual directional hash chains (DDHC) for the Logical Key Hierarchy (LKH) rekeying protocol. This enables a member that has missed up to m consecutive key updates to recover the missing decryption keys without asking the group controller key server (GCKS) for retransmission. Conducted simulations show considerable improvements in the ratio of decrypted messages and in the rekey message overhead in high packet loss environments.

Keywords: Secure multicast, Logical key hierarchy, Group rekeying protocol, Group key management, Dual directional hash chains, Multicast recovery scheme, Secure group communications, Group controller key server, Mobile applications, High packet loss

5.1 Introduction

Fixed-mobile convergence in 4G networks will lead the way into more complex and resource-hungry mobile applications such as mobile TV, video teleconferencing and stock quote distribution. Network operators and service providers face an important dilemma: on one hand, the demand for such applications is strong, but on the other, the high bandwidth consumption in the radio access network dramatically increases the price for such services, thus making them unattractive. Multicast transmission is a key factor in the successful deployment of high bandwidth applications offered in both fixed and mobile networks. In order for the operators to charge customers for their requested services, multicast security mechanisms must be put in place to provide key distribution, data origin authentication, and policy management.

To ensure data confidentiality of a multicast flow, all members of a multicast group share the same key (referred as the group key) that is used for encrypting the data. The group key management (GKM) schemes are responsible of generating and updating the keys to ensure the forward and backward secrecy. Forward secrecy is an important property which makes it impossible for a revoked or a departed member to decrypt the multicast data or rekey messages after leaving. Similarly, the backward secrecy property ensures that the multicast messages prior of a member joining the group remain undecryptable for that member. Another very important security property is its resiliency to collusion attacks. A revoked member could exchange all its keys with another revoked user to decrypt messages they were not entitled to. However, a revoked member could also collude with a current member who knows the actual group key to learn group keys that were used after the former was evicted but before the latter joins the group.

Most group key management protocols are tree-based, meaning that the Group Key Controller Server (GCKS) constructs a tree where each logical node possesses a key shared among a restricted number of current members of a multicast group. Among the many group rekeying protocols that have been proposed, each having their own merits, the logical key hierarchy (LKH) and the subset-difference revocation (SDR) algorithms are the most two popular. The former is a stateful protocol in which the rekeying overhead is strictly correlated to the state of the membership (the logical tree) during the rekeying instance whereas the rekeying overhead of SDR depends on the membership (tree with subsets) over the entire multicast session. LKH seems to outperform SDR in immediate rekeying or small batch rekeying Weifeng Chen (2008); Zhu *et al.* (2003). In fact, the rekeying overhead in LKH is fairly stable whereas it increases parabolically in SDR with the increasing number of revoked users Zhu et Jajodia (2003); Ioannidis *et al.* (2005). Thus, LKH offers more scalability and stability over time and

is the most suitable rekeying protocol for generic multicast services in 4G networks Chen et Dondeti (2003b).

Other approaches for scalable rekeying such as one-way function trees (OFT) Sherman et McGrew (2003) and ELK Perrig *et al.* (2001) also involve the use of a hierarchical key tree in which keys at higher levels of the tree are needed by more members than keys at lower levels. More recently, Maximum Distance Separable (MDS) Raj et Lalith (2009) codes have been introduced as an alternative to encryption algorithms in tree-based group key management protocols where members of a multicast group are able to recover the group shared key through the erasure decoding of MDS codes.

A key distribution that is reliable or better yet, offers self-healing properties is of particular interest to mobile environments where users can experience high packet loss. Stateless group management protocols have a clear advantage by giving a legitimate user the ability to extract the new group key despite the number of missed rekeying materials. A self-healing mechanism for SDR enabling members to reconstruct missed group keys has also been proposed Zhu *et al.* (2003). On the other hand, LKH lacks robustness against packet loss and makes it impossible for a member who missed a single rekey message to decrypt any subsequent rekey or data messages.

The main contribution of this paper is to present two self-healing schemes for LKH based on the dual directional key chains (DDHC) and to show how well it reacts to different mobility scenarios in 4G networks. The remainder of this document is organized as follows. Section 5.2 discusses the existing works in more details followed by the proposed self-healing schemes for LKH in section 5.3. An analytical analysis and the simulation results are detailed in sections 5.4 and 5.5 respectively. Finally, a brief summary outlining key observations concludes the paper.

5.2 Background concepts and related work

5.2.1 Logical Key Hierarchy group rekeying protocol

The simplest approach for group rekeying is for the group key server to individually encrypt the group key with the shared private key of each member and unicast it. Therefore, the more members are part of a multicast group, the more rekey messages must be sent unicast. Obviously, such method is not scalable since the rekeying cost increases linearly with the group size.

Tree-based multicast key distribution D. *et al.* (1999); Wong *et al.* (2000) considerably helps to reduce the bandwidth overhead. In this rekeying scheme, the GCKS constructs and maintains a tree where the leaves are the members of a multicast group. Each node of the tree

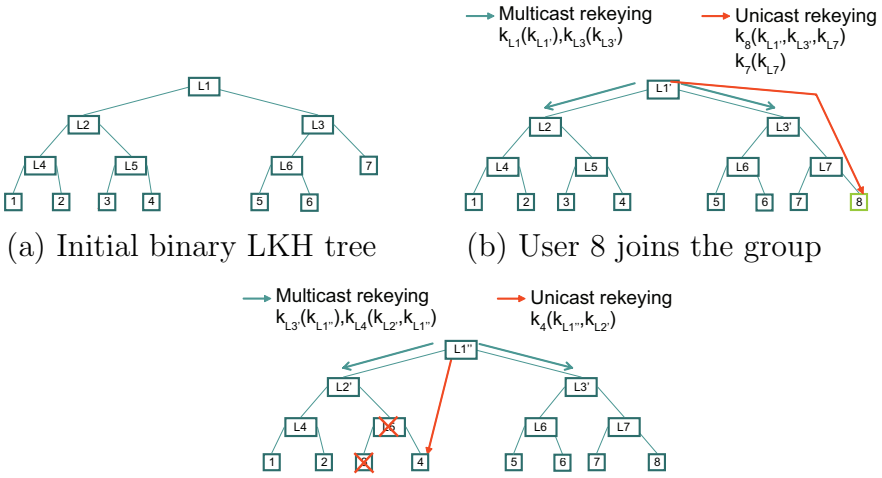
is associated with key. The root key is shared among all multicast members and is therefore the group key. The other keys are used to encrypt other keys during a rekeying process and are known as key encryption keys (KEKs). Every member is aware of the keys along the path from the leaf to the root.

When a new member joins the group, the GCKS adds the leaf node to its logical tree. To satisfy the backward secrecy, all the keys along the path from the corresponding leaf node to the root need to be refreshed. The GCKS then encrypts these new keys with the private key of the new node and sends it unicast. Also, the GCKS individually encrypts these new keys using their previous key and multicasts it for every existing members to decrypt the rekey message and update its corresponding key. Because the new member ignores the old previously used group and key encryption keys, the backward secrecy is respected.

When a node is revoked, the GCKS refreshes all the keys known by the revoked member and then deletes the corresponding leaf node. Each of the new key is individually encrypted using the keys from the siblings of all the logical nodes along the path from the root to the deleted leaf node. Because those encryption keys are unknown to the departed member, the forward secrecy is assured.

Figure 5.1a illustrates an example of a balanced binary logical tree handling a small multicast group of 7 members (named 1 through 7). With the arrival of a new authorized member (8) as showed in figure 5.1b, the GCKS adds a leaf node next to 7 and updates the keys from the path linking the new node, or its sibling 7, to the root. The multicast message is constructed by encryption each new key with the previous KEK, resulting in the following rekey message $K_7(K_{7\_8}), K_{5\_7}(K_{5\_8}), K_{1\_7}(K_{1\_8})$, where $K_2(K_1)$ means key K_1 is encrypted using key K_2. The unicast message sent to 8 is simply the concatenation of all the new keys encrypted with the private shared key of 8: $K_8(K_{7\_8}, K_{5\_8}, K_{1\_8})$. As depicted in figure 5.1c, the revocation of user m3 will force the GCKS to update all the keys the departed member possesses and send the following rekey multicast and unicast message respectively: $K_{5\_8}(K'_{1\_8}), K_{1_2}(K'_{1\_4}, K'_{1\_8})$ and $K_4(K'_{1\_4}, K'_{1\_8})$.

Stateful key distribution performance is extremely dependant on the rate loss of rekey messages. As depicted in figure 5.2, when a member j misses a rekey message, it may be unable to decrypt subsequent rekey and data messages. More precisely, in a rekey event caused by a join event, only the keys from levels above the missed rekey message can be decrypted. On the other hand, in a leave event, the rekey message subsequent to a missed one can be decrypted only if its level is greater or equal to the level of the missed rekey message.



(a) Initial binary LKH tree (b) User 8 joins the group

(c) User 3 leaves the group

Figure 5.1 LKH execution example

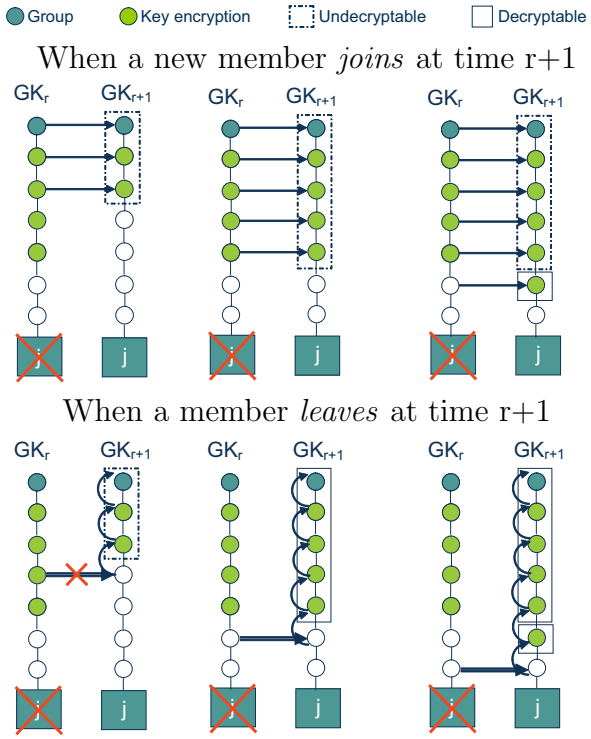


Figure 5.2 Detailed LKH key manipulation for join and leave events

5.2.2 Optimized key recovery mechanism for LKH

A member who is unsuccessful in decrypting keys can explicitely request the GCKS for retransmission. A simple but inefficient way is for the GCKS to send the most current keys of all the nodes in the path from the root to the leaf node corresponding to the node who requested a key recovery. Such technique may make the GCKS send keys the member already possesses. The chamois key recovery scheme Cho *et al.* (2004) objective is to enable the recovery of any group key and only the useful KEKs by simply keeping the current key-tree along with the following information that reflects the update history of each node:

- an array of w 2-bit flags, where w is the maximum amount of group data and rekey messages a node can buffer, indicating if event i is a join or leave event
- the last index of an event that causes an update of the corresponding key

Also, the GCKS generates the group key by computing $GK_i = PRF(sKey, i)$ where PRF is a one way pseudo-random function, sKey is secret key only known to the GCKS and i is the index of the event. By using this information, the GCKS can compute the level of missed rekey message and thus effectively send only the required keys to the member who requested it.

5.2.3 Reliable Key distribution

Reliable key distribution schemes essentially aim at a better reception rate of rekey messages in soft real-time at the expense of higher bandwidth overhead. By assigning weights to nodes, sending hints or simply resending the keys multiple times before the next event that triggers another rekey message, the number of undecryptable data messages will be considerably decreased.

In Zhang *et al.* (2003), authors have proposed Proactive FEC in which, instead of resending the rekey messages, the GCKS uses a Reed-Solomon erasure (RSE) coder on blocks of k encrypted keys to generate h forward erasure encoding (FEC) redundant information also known as parity packets. By multicasting blocks of the encrypted keys and parity packets, members may be able to recover lost encrypted keys. The number of parity packets is computed from the proactive factor defined as $(h + k)/k$. The proactive factor must be properly adjusted in order to limit the bandwidth overhead and avoid sending too many parity packets to nodes who are not experiencing a considerable packet-loss rate.

Authors in Setia *et al.* (2002) have proposed the weight key assignment (WKA) and the batched key retransmission (BKR) algorithms. The first aims at assigning replication weights to encrypted keys based on the members' loss rates. The GCKS then packs the encrypted keys with similar weights into the same set of packets. The higher the weight on a key is,

the more frequent it will be retransmitted. BKR aims at packing the keys needed by several members instead of processing them one by one. WKA-BKR has been shown to have a lower bandwidth overhead than Proactive FEC scheme over a wide range of group sizes and membership dynamics and in network loss conditions. However, because RSE encoding used in FEC is more efficient than the simple key retransmissions, Proactive FEC has a lower latency.

An hybrid approach of the two previous reliable key distribution methods in which the authors essentially take the WKA-BKR algorithm and replace the key retransmissions with FEC (parity) blocks. WFEC-BKR Zhu *et al.* (2003) thus benefits from a low bandwidth overhead and a relatively low latency.

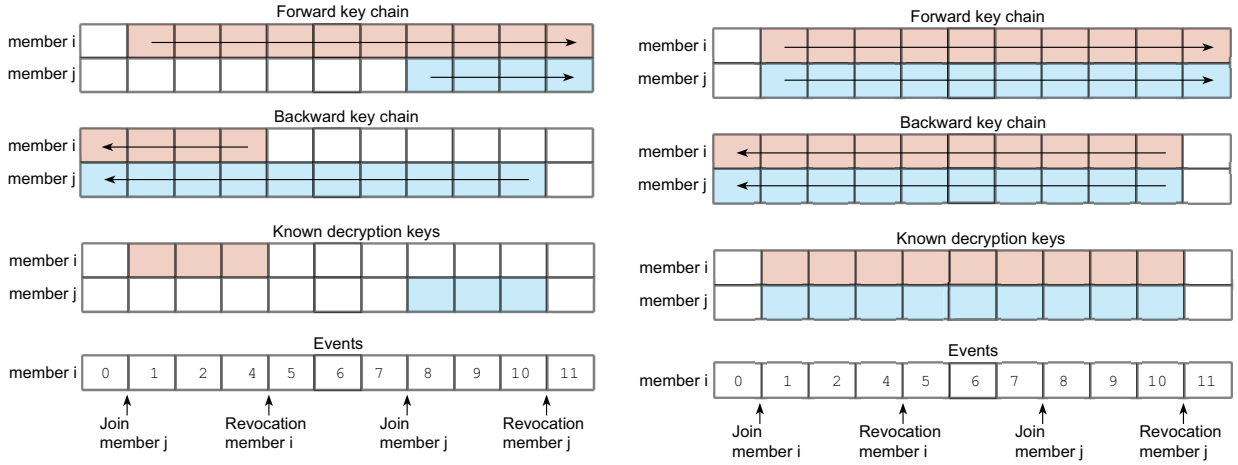
5.2.4 Self-Healing key distribution

All the reliable key distribution methods described above increase the reception rate of rekey message at the expense of a higher overhead. Self-healing key distributions enable a member to recover a key that has been lost without sending a retransmission request and without applying the reliable key distributions mentioned above.

Unconditionally self-healing secure schemes using threshold access structure based on polynomial interpolation, more commonly known as Shamir's secret, have been the first to be proposed. The pioneers in Staddon *et al.* (2002) have provided formal definitions, lower bounds on the resources as well as some constructions of unconditional self-healing key distribution schemes based on polynomial functions. However, their solution suffers from inconsistent robustness, high overhead and expensive maintenance cost. By generalizing the definitions and lowering the bounds, authors in Blundo *et al.* (2007); Liu *et al.* (2003); Hong et Kang (2005) proposed some more efficient constructions. The use of a sliding window in More *et al.* (2003) makes error recovery consistently robust while the reuse of masking polynomials reduces broadcast size and key storage significantly. However, because self-healing schemes based on Shamir's secret limit the number of revoked users to the degree of the polynomial, these solutions are unattractive. Using a more generalized vector space secret Padro *et al.* (1999); Tian *et al.* (2008) by considering a monotone decreasing family of rejected subset of users instead of a monotone decreasing threshold structure helps solve this issue.

By slightly relaxing security properties, computationally self-healing secure schemes Dutta *et al.* (2008, 2007); Kausar *et al.* (2007) are much more efficient. In Shi *et al.* (2007), a time-limited node revocation based on Dual Directional Hash Chains (DDHC) has been proposed for securing multicast flows in wireless sensor networks (WSN). A DDHC consists of a forward and backward key chains going in opposite directions and generated by repeatedly applying a one way hash function from an initial random seed. When a member joins a group, the

GCKS securely sends it the current forward key and the backward key corresponding to its revocation time. Upon reception, the member constructs its backward chain by applying the one-way function on the received backward key. As shown in figure 5.3(a), because the group traffic key is generated by combining both the current forward and backward key, the legitimate member possesses all the necessary keys for decryption between its joining and revocation time. Note however, that the member is unable to compute the forward keys prior to its joining time or the backwards keys after its revocation time, thus respecting the forward and backward secrecy. This self-healing construction is very efficient but suffers from a serious security drawback in which two revoked nodes can collude to retrieve decryption keys they were not entitled to, or more precisely between the time from the first revocation and the second join as shown in figure 5.3(b). The issue exists for the same node who wants to rejoin the same group after being revoked earlier. It is also important to note that this solution was developed in the context of a single WSN where their security and traffic requirements greatly differ from managing multimedia streams in FMC scenarios. While the proposed solution makes use of DDHC to provide self-healing properties, it offers a much more global approach by integrating a hierarchical architecture to improve scalability in a context of supplying different media multicast traffic in multiples fixed and mobile networks where security policies differ from one administrative domain to another.



(a) non collusion scenario (b) collusion scenario

Figure 5.3 DDHC keys in a non collusion and collusion scenarios

Authors in Zhu *et al.* (2003) proposed a group key recovery scheme that adds an m-recoverability self-healing property to SDR in which the maximum number of previous group keys a legitimate user can recover is m. As per the SDR concept, member nodes are partitioned into m+1 subgroups depending upon their membership duration. For each group

rekeying, the GCKS generates a one-way key chain of size m+1 (i.e., for a rekeying at time T(i), the chain would be $K^m(i), K^{m-1}(i), ..., K^0(i)$ where $K^0(i) = H(K^1(i)) = H^2(K^2(i)) = ... = H^m(K^m(i))$ and H is a one-way hash function). Each key of the chain is dedicated to a specific subgroup according to the time the members joined the group. The longer a member has been in the group, the higher the degree of the key (up to m) it receives with which it can derive the keys up to $K^0(i)$. The GCKS securely and reliably sends these keys by encrypting them with their corresponding subset key. The final step is to multicast the group key securely. To address the colluding attack between a revoked node and a newly joined member, the group keys are encrypted with a combination (using a XOR) of a previously received key and the current one from the chain. More precisely, for a current member that joined at T(j), $(i-m) < j < i$, it receives $K^{i-j}(i)$ which enables to recover the keys between $K^0(j)$ and $K^0(i)$ by decrypting i-j keys of the m keys from the multicasted rekey message by the GCKS. This solution unfortunately inherits the main SDR drawback of increasing communication complexity when the number of revoked users grows. In fact, it is tightly related to the number of subsets in the multicast group, and therefore, the storage, computation and bandwidth overheads can be important drawbacks.

5.3 Self-healing schemes for LKH

This section presents the proposed self-healing schemes for LKH based on the DDHC. Both schemes allow members to reconstruct the decryption keys without asking the GCKS if it missed up to m consecutive rekey messages. In scheme I, members profit from the self-healing properties in both the join and leave events, but it is vulnerable to rejoining/colluding attacks and revocation processes are more complex to manage. On the other hand, scheme II temporary disables self-healing on a leave event, but is much more robust and scalable.

5.3.1 Definitions and notations

The following definitions and notations (table 1) will be used in the description of the schemes:

- Self-healing period : maximum number of consecutive rekey messages a member node can miss before being unable to regenerate the keys;
- Refreshing period: the window of time before a new set of forward and backward keys are generated for the LKH hierarchy (refreshing period $\geq$ self-healing period, typically refreshing period >> self-healing period);
- Batch revocation period: time between two batched revocation events;
- Logical node: a node in the LKH tree that is not a leaf nor representing a participant

in the multicast group;
- Group members: participants in the multicast group (correspond to the leaf nodes in the LKH tree);
- Level in the LKH tree: level or depth in which the logical node is located in the LKH tree; the higher the level of a logical node, the more child nodes it possesses (the root node level is 1).

Table 5.1 Sets, variables and notations

I	Set of all logical nodes in the LKH tree
J	Set of all the group members in the multicast group
S_i	Set of all siblings of $i \in I$ ($S_i \subset I$)
NP_j	Set of logical nodes $i \in I$ forming the path between the member j and the root logical node of the LKH tree ($NP_j \subset I$)
f_i^k	k-ith forward key of $i \in I$
b_i^k	k-ith backward key of $i \in I$
sf_i	Forward key seed of $i \in I$
bf_i	Backward key seed of $i \in I$
PK_j	Private key for member $j \in J$
$rekey^t$	Multicast rekey message triggered at $event^t$
$event^t$	The event that occurred at time t (join or leave)
k_i^t	Position of the pointer of the backward and forward key chain of $i \in I$ at $event^t$
m_i	Self-healing period of $i \in I$ ($m_i \geq 0$)
l_i	Level of the logical node $i \in I$ in the LKH tree. Note that the level of a node is larger than that of the parent node by 1 and the root node level is 1.
$l_j^{rekey^t}$	Highest level of the logical node $i \in NP_j$ of the updated keys for member $j \in J$ in $rekey^t$. For example, in figure 5.4, for $event^r = join$ of member 8, $l_1^{rekey^r} = 1$ and $l_7^{rekey^r} = 3$.

5.3.2 Scheme I

The core idea in the proposed schemes is the construction of a DDHC for every logical nodes in the LKH tree. More precisely, the GCKS constructs backward key chains for each level of the LKH tree which must be long enough to handle the join and leave events that occur during the refreshing period. When the GCKS updates the keys at $event^t = join/leave$ of member $j \in J$, it increments $k_i^t \forall i \in NP_j$ and updates the forward key by applying once the one way function $f_i^{k_i^t} = H(f_i^{k_i^t - 1})$ and deletes the old backward key $b_i^{k_i^t - 2}$ from the backward chain. The current encryption/decryption key is simply a combination (such as using a XOR)

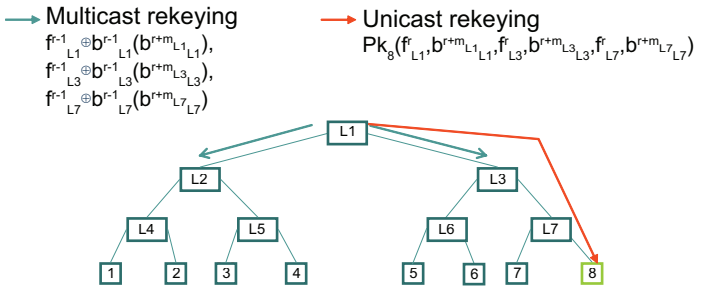
of both the forward and backward key at the pointer current position k_i. The GCKS may also update m_i to adjust the self-healing period according to the key loss rate.

In more details, when $event^r = join$ of member j, the GCKS provides the triplet $PK_j(f_i^{k_i^r}, b_i^{k_i^r+m_i}, m_i) \forall i \in NP_j$ to the newly joined member j. For the current subscribed members, only the backward key $f_i^{k_i^r-1} \oplus b_i^{k_i^r-1}(b_i^{k_i^r+m_i}, m_i) \forall i \in NP_j$ is sent.

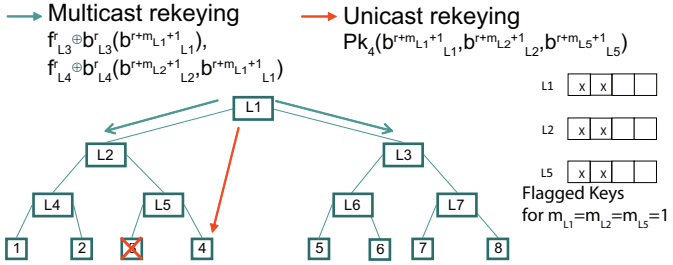
When a member j is revoked, the GCKS must make sure to send the updated keys to all subscribed members without using an encryption key known by the revoked member. Because $m_i \forall i \in NP_j$ keys were sent in advance by the GCKS, the complexity in handling secure rekey transmission is a lot higher than the regular LKH process. In fact, after a member is revoked at time/event r, it still possesses the valid decryption keys $f_i^{k_i^t} \oplus b_i^{k_i^t} \forall i \in NP_j, r-1 \leq t \leq r-1+m_i, m_i \geq 0$ which enables it to decrypt $m_i$data messages and rekey messages after the revocation rekeying process has been completed. To address this issue, upon a revocation of member j, the GCKS must flag every key from every chain for all $I \in NPj$ the revoked member still possesses. When the GCKS sends a subsequent updated key, it must ensure that the encryption key that is being used is not flagged. In such cases, the GCKS searches a logical node from a higher level with a current unflagged backward and forward key until reaching the member node in which case the key will be sent in unicast (see figure 5.4 for an example). By doing so, the rekey message will contain more keys than the original LKH scheme because every time the GCKS goes up 1 level in the LKH tree, the number of keys to send for a rekey event is multiplied by the degree of the LKH tree. Thus, in an environment with frequent revocation, this scheme is unsuitable because it dramatically increases the overhead.

To limit the overhead, it is important for this scheme to find the optimal self-healing period per logical node in the LKH tree. A high $m_i, i \in I$ helps members recover keys (lower key loss rate), but also increases the overhead during the rekey messages following a revocation. Based upon the number of retransmission request, the GCKS must find the best tradeoff between the self-healing period and the performance degradation.

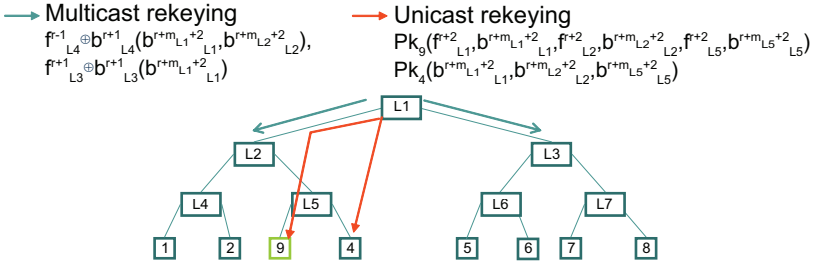
Upon receiving the rekey message, the group members repeatedly applies the hash function to the new backward key until resulting in the current backward key. The same number of iteration to reach the current backward key from the new one will be applied to the current forward key to update its forward key. The new member j, on the other hand, constructs its initial backward key chain by simply applying the one way hashing function the number of times defined by $m_i, \forall i \in NP_j$to the received backward and forward key. Figure 5.5 illustrates an example how a member who missed a rekey message decrypts and regenerates the missed keys.



(a) Event r : User 8 joins



(b) Event r+1 : User 3 leaves



(c) Event r+2 : User 9 joins

Figure 5.4 Self-healing LKH key distribution example with join and leave events for scheme I

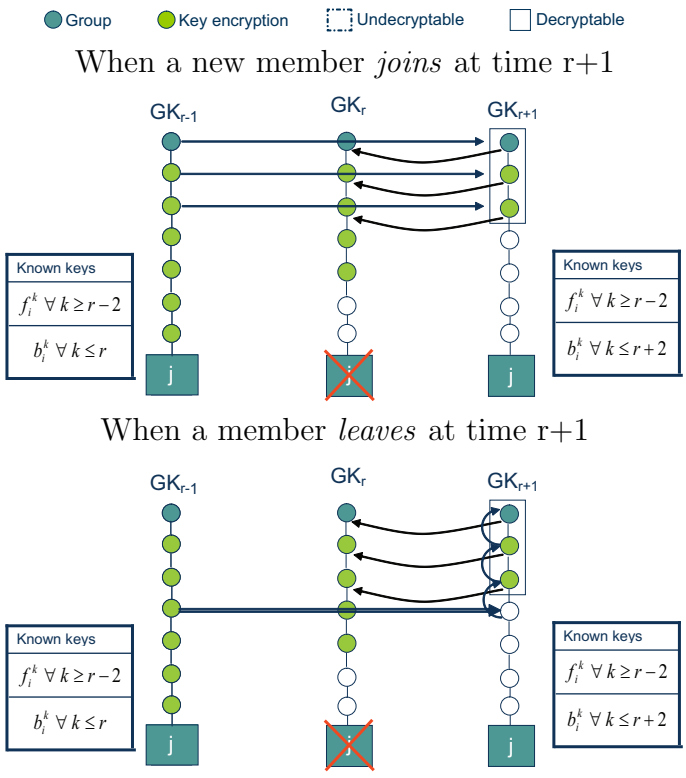


Figure 5.5 Example of a self-healing key recovery

5.3.3 Scheme II

The idea behind this second scheme is to find a tradeoff between keeping the self-healing properties for every events and its performance degradation and vulnerability to collusion attacks. The objective is to address rekeying cost increase and the colluding attack issues without so much affecting the self-healing properties.

The rekeying cost increases in the first scheme due to the valid $m_i$decryption keys $f_i^{k_i^t} \oplus b_i^{k_i^t} \forall i \in NP_j, r-1 \leq t \leq r-1+m_i, m_i \geq 0$ known by member j after $event^r = leave$, forcing the GCKS to encrypt the same key multiple times using keys from higher-level siblings not connected to the revoked member. Therefore, if $event^r = leave$ of member j, there are two ways that can handle this issue:

1. The GCKS securely sends $b_i^{k_i^r+m_i+m_i^{'}}, m_i^{'} \forall i \in NP_j$ where m_i is the self-healing period of $event^{r-1}$ and use $b_i^{k_i^r+m_i}$ for the encryption key of the next event;
2. The GCKS reconstructs the backward or the forward key chain after each user revocation.

These techniques provide immediate revocation but also temporarily disable the self-healing property. Thus, the backward keys sent after a member is revoked must be sent reliably

in order for subscribed members to decrypt the next data and rekeying messages. Because frequent revocation events is problematic, batch revocation must be used with a timeout that offers the best tradeoff between the key loss ratio and an acceptable user revocation delay is used.

The collusion attacks in scheme I are made possible because of the reuse of the same backward and forward key chains after a user is revoked. The simplest way to solve both problems is for the GCKS to perform one of the two following tasks after a revocation event:

1. reconstruct the backward key chain according to the estimated number of join events between two batched revocation events;
2. generate a new forward key.

Note that both strategies can simultaneously be performed if the security policies forbidden any valid encryption information to be known by a revoked member. However 2 keys must be sent instead of one for each legitimate node, thus increasing the bandwidth complexity.

The recommended strategy to address the collusion attacks and the increase in communication complexity after a member revocation is to reconstruct a new backward key chain between batched revocation events. Because the chain must only accommodate the new join events between 2 batched revocation events, it greatly reduces the storage requirements. At each batched revocation timeout, the GCKS securely and reliably sends $b_i^{m_i}, m_i \forall i \in NP_j$ of each revoked members j.

5.4 Analytical analysis

5.4.1 Security observations and proofs

Scheme I incurs the following security observations:

1) Scheme I offers self-healing regenerating properties without sending a request to the GCKS as long as the encryption key has not been missed more than m consecutive times. Consequently, LKH now has the m-statelessness property which guarantees that a member can go offline and miss as much as m rekey messages and still be able to participate in the multicast group upon the reception of the rekey message on his return.
   (a) A member can miss even more than m consecutive rekey messages and still be able to decrypt some or all of the key encryption keys (see figure 5.6 for an example).

Proof 1. A member who successfully decrypted a backward key i from a rekey message for $event^r$ possesses valid keys for $b_i^1, ..., b_i^{k_i^r+m_i} \forall i \in NP_j$. Therefore, having missed m_i consecutive key updates, a member receives the rekey message $f_i^{k_i^r+m_i} \oplus b_i^{k_i^r+m_i}(b_i^{k_i^r+m_i+1+m_i^{'}}, m_i^{'}) \forall i \in NP_j$ for $event^{r+m_i+1} = join/leave$ where $m_i^{'}$ is the new self-healing period. The decryption

key $f_i^{k_i^r+m_i} \oplus b_i^{k_i^r+m_i}$ for $event^{r+m_i+1} = join/leave$ is known by the member and can therefore reconstruct the backward chain $H^{m_i^{'}}(b_i^{k_i^r+m_i+1+m_i^{'}})$, $H^{m_i^{'}-1}(b_i^{k_i^r+m_i+1+m_i^{'}}), ..., b_i^{k_i^r+m_i+1+m_i^{'}})$ and update the forward key $f_i^{k_i^r+m_i} = H^{m_i}(f_i^{k_i^r})$. □Proof 1.a. A member may miss more than m rekey messages and still be able to decrypt the rekey messages and reconstruct the chains because:

- Consecutive rekey messages do not necessarily contain updated keys from the same logical nodes (this is especially true for high level nodes in the LKH tree). Note however that the group key used to encrypt data is the lowest level key of the LKH tree and therefore always updated during an event.
- Updated keys issued from a member revocation or following a member revocation (see figure 5.6) are encrypted with a higher level KEK which may have not been updated in every of the last m events.

□

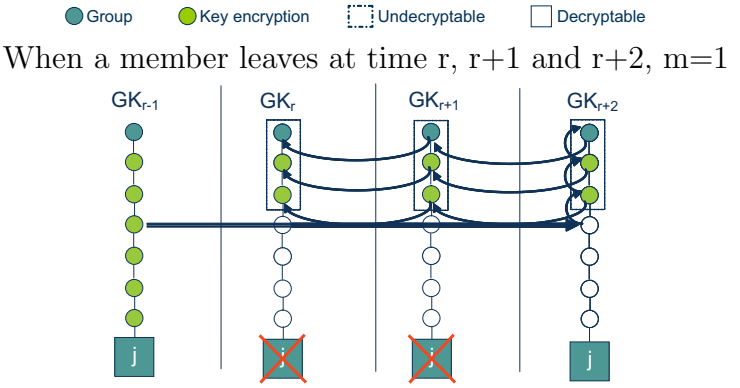


Figure 5.6 Example of key recovery for scheme I after missing more than m rekey messages

2) Inversely, a member is unable to decrypt a key in $rekey^t$ when the decryption key has been missed more than m consecutive times. More specifically, if $rekey^t \forall r \leq t \leq r + m_i, i \in NP_k$ messages have been missed by member k, then if $event^{r+m_i} = join$, then all decryption keys for levels $min(l^j_{rekey^t}) \forall r \leq t \leq r + m_i$ for members $i \in J$ are invalid. On the other hand, if $event^{r+m_i} = leave$, then all decryption keys for levels $min(l^j_{rekey^t}) \forall r \leq t \leq r + m_i$ are invalid if $min(l^j_{rekey^t}) \geq l^j_{rekey^{r+m_i}} \forall r \leq t \leq r + m_i$.

Proof 2. A backward key $b_i^{k_i^r+m_i}$ is encrypted with $f_k^{k_i^r-1} \oplus b_k^{k_i^r-1}, l_k \geq l_i$. More precisely, if $event^r = join$,$l_k = l_i$, else if $event^r = leave$ then $l_k > l_i$. After m+1 events, keys will be encrypted with keys $f_k^{k_i^r+m_i+q} \oplus b_k^{k_i^r+m_i+q} \forall q \geq 1, l_k \geq l_i$ which are unknown for members who missed m+1 or more consecutive updates of the same key. In $rekey^t$, each key sequentially corresponds to a level in the LKH tree going from $0, 1, ...l^j_{rekey^t}$. Therefore, if $b_i^{k_i^r} \forall i \in NP_j$

is present in $rekey^t$, then $b_k^{k_i^r} \forall i \in NP_j, 1 \leq l_k \leq l_i$ keys are present as well. Consequently, if a member j missed m+1 consecutive rekey messages, it would have missed m+1 updates of keys of logical nodes i where $1 \leq l_i \leq min(l^j_{rekey^t}) \forall r \leq t \leq r + m_i$and be unsuccessful in decrypting any key or data encrypted with it. □

3) Scheme I is computationally secure and provides forward and backward secrecy.

Proof 3: Because it is computationally hard to inverse a one-way hashing function $(H^{-1}(b_i^k))$ and thus making it impossible for a member who possesses keys $H^j(key) \forall j > c$ to find $H^c(key) \forall C \geq 0$. More specifically, the backward key chain is constructed as follows $H^i(sb_i)$, $..., sb_i \forall i \geq 1$ allowing a member who possesses $H^j(key)$ to compute all the prior keys but not the next keys, thus assuring the forward secrecy. On the other hand, the forward key chain $(sf_i, ..., H^i(sf_i) \forall i \geq 1)$ makes it impossible for a member who possesses keys $H^j(key)$ to compute all the next keys but not the prior keys , thus assuring the backward secrecy. Combining the forward and backward key thus provides forward and backward secrecy. □

4) Scheme I is vulnerable to rejoining/colluding attacks.

Proof 4: This scheme simply applies DDHC for each level of the LKH tree and thus inherits the same drawbacks as DDHC. The proof is trivial by extending the issue illustrated in figure 5.3(b) for every level of the LKH tree. □

5) A revoked member will be able to continue to decrypt data messages for m events.

Proof 5: After $event^r = leave$, the GCKS encrypts the data with $f_1^{k_1^r} \oplus b_1^{k_1^r}$. But because it has sent $b_1^{k_1^r + m_1 - 1}, m_1 \geq 1$to the revoked member at the previous event, the revoked member knows $b_1^{k_1^r}, ..., b_1^{k_1^r + m_1 - 1}$ which enables it to decrypt the data messages for up to m_1 events. □

Scheme II incurs the following security observations:

1) The m-statelessness property of scheme I is still preserved, however it is only effective as long as the key is not refreshed by a member revocation event.

Proof 1. The previous proof 1 still stands as long as the m consecutive missed rekey messages are generated from $event^r = join \forall r \leq t \leq r + m_i$. In fact, when $event^r = join$ for member j, the GCKS sends $f_i^{k_i^r - 1} \oplus b_i^{k_i^r - 1}(b_i^{k_i^r + m_i}) \forall i \in NP_j$. A member who received a previous rekey message from $event^t = join \forall r - 1 - m_i^{'} \leq t \leq r - 1$, is therefore capable of decrypting $rekey^r$. When $event^r = leave$, the GCKS reconstructs a new backward chain and sends $f_i^{k_i^r - 1} \oplus b_i^{k_i^r - 1}(b_k^{m_k^{'}}) \forall i \in S_{NP_j}, k \in NP_j$ which not only requires the reception the last updated key for i, but it is crucial for members to receive the new backward key to construct the new chain and re-establish the self-healing property. □

2) Observation 2 from scheme I still stands as long as the m+1 consecutive missed rekey messages are sent from join events. If $event^r = leave$, then decryption keys $f_i^{k_i^t} \oplus b_i^{k_i^t} \forall i \in NP_j, t < r$ are invalid.

Proof 2. The proof 2 from scheme I still stands for rekey messages issued from $event^t = join$. When $event^t = join$, the GCKS constructs a new backward key chain which must be used immediately to decrypt the next data/rekey messages for the logical nodes in the path from the revoked member to the root. □

3) Observation 3 from scheme I and its proof stands.

4) Scheme II is no longer vulnerable to rejoining/colluding attacks.

Proof 4: When a member j is revoked at $event^r$, the GCKS computes a new backward key chain making all past and current decryption keys invalid for the affected logical nodes. A member must receive $rekey^r$ in order to decrypt future data and rekey messages. When member j rejoins the same group at $event^k \forall k > r$, it will receive $PK_j(f_i^{k_i^k}, b'^{k_i^k+m_i}_i, m_i) \forall i \in NP_j$ where $H^i(b'^{k_i^k+m_i}_i) \neq b_i^{k_i^t} \forall i \geq 0, t < r$. In fact, the keys between the member revocation and its second rejoin events cannot be recomputed because of an unknown information (the new backward key chain), thus making the solution resilient to rejoining attacks. The same missing information occurs when two revoked members (or one revoked and one recently joined member) share their backward and forward keys to compute decryption keys they were not untitled to receive (keys between the earliest revocation and the latest join events). □

5) A revoked member will immediately be revoked no longer be able to continue decrypt data messages for m events.

Proof 5: A $event^r = leave$ of member j triggers the GCKS to generate a new backward key chain $b_i^{k_i^r} \forall i \in NP_j$ which makes the new encryption key not entirely based on previous known information. Because the root is always part of NP_j and the new information will only be shared among the subscribed members, it disables the ability for the revoked member to decrypt future data messages. □

5.4.2 Efficiency Analysis

This subsection compares the key storage and rekeying cost of our schemes with LKH.

Member and GCKS key storage

Table 5.2 presents a summary the key storage comparison for the worse case scenarios. Because $L^I \gg L^{II} \geq 0$, where L^I is the length of the backward key chain for scheme I is very memory expensive for the GCKS compared to the others. In fact, it must construct the backward key chain for the complete refreshing period as opposed to scheme II where the backward key chain can be much shorter to supply the number of join events only for the

batched revocation period. The impact is limited on the members if the self-healing period is kept to a minimum but still accommodating its rate loss.

Table 5.2 Worse case key storage comparison

	LKH	Scheme I	Scheme II
GCKS storage	$\frac{d \cdot n-1}{d-1}$	$\frac{d \cdot n-1}{d-1} \times (L^I + 1) - n \cdot L^I$	$\frac{d \cdot n-1}{d-1} \times (L^{II} + 1) - n \cdot L^{II}$
Member storage	$\log_d n + 1$	$(max(m_i) + 1) \times \log_d n + 1$	$(max(m_i) + 1) \times \log_d n + 1$

Rekeying cost

Another important evaluation metric is the overhead when processing a rekeying triggered by a join or leave event. As shown in table 5.3, the number of keys to send when a new member joins the group is higher for the proposed self-healing schemes because the decryption key is composed of a forward and a backward key whereas the decryption key is a single randomly chosen key in LKH. Note that the numbers of keys to send in scheme I depends on the number of revocation members in the last m events and their location in the LKH tree. The best case occurs when they were no revocation in the last m events (and therefore no flagged keys) where the overhead is the same as LKH for $event^r = leave$ and $\log_d n$ keys bigger for $event^r = join$ because 2 keys instead of 1 must be sent to the new member. On the contrary, the worse case is when at least 1 member per group of d members sharing the same lowest level logical node of the tree is revoked in the last m events, thus with all the keys flagged by the GCKS. Such extreme case leads to a number of keys that increases exponentially, but the probability of occurrence is relatively low.

Table 5.3 Rekeying cost comparison

	LKH	Scheme I	Scheme II
New key distribution amount triggered by a join request	$2 \cdot \log_d n$	$[3 \cdot \log_d n, \log_d n + \frac{d}{d-1} \times (d^{\log_d n} - 1)]$	$3 \cdot \log_d n$
New key distribution amount triggered by a leave request	$d \cdot \log_d n$	$[d \cdot \log_d n, -\log_d n + \frac{d}{d-1} \times (d^{\log_d n} - 1)]$]	$d \cdot \log_d n$

5.5 Experimental results

This section compares LKH with the proposed self-healing addon schemes by showing empiric results obtained using Qualnet 4.5.1 simulator from Scalable Networks. The scenario details and along with the results follow.

5.5.1 Performance metrics and primary factors

The scenario details are described in table 5.4.

Table 5.4 Experiment details

Static factor	Description	
Simulation Time	20 minutes	
Number of nodes	100	
Terrain area	22km x 8km	
Number of executions per scenario	10 different seeds	
Multicast application protocol	MCBR	
Throughput	1Mbps	
Base Station properties	Wireless Connection Type Number Radio Range	WiMAX 3 8km
Node position	Random	
Node direction	Random way-point	
LKH tree	Balanced binary tree	
batch revocation timeout	5 minutes	

Conducted simulations are performed in regards to the ratio of undecryptable data messages and the rekeying cost with "a one factor at a time" experiment. The ratio of undecryptable data messages is defined as the number of data messages that have unsuccessfully been decrypted on the total number of received data messages. Therefore, the loss of data messages is not considered in the experiment. The rekeying cost is the bandwidth overhead quantified with the number of sent keys caused by a join/leave event or a request from a legitimate member. Table 5.5 shows the primary factors considered for each session.

5.5.2 Empirical results and analysis

Figure 5.7a shows that member's mobility negatively impacts LKH performance while its effects are contained in the proposed self-healing schemes. It is safe to assume that the higher

Table 5.5 Executions details

Impact of	Primary Factors	
mobility	Rekey Retransmissions Buffer Revocation	0 0s None
retransmissions	Mobility Buffer Revocation	30 mps 0s None
buffer	Rekey Retransmissions Mobility Revocation	3 10 mps None
revocation	Rekey Retransmissions Buffer Mobility	0 0s 30 mps

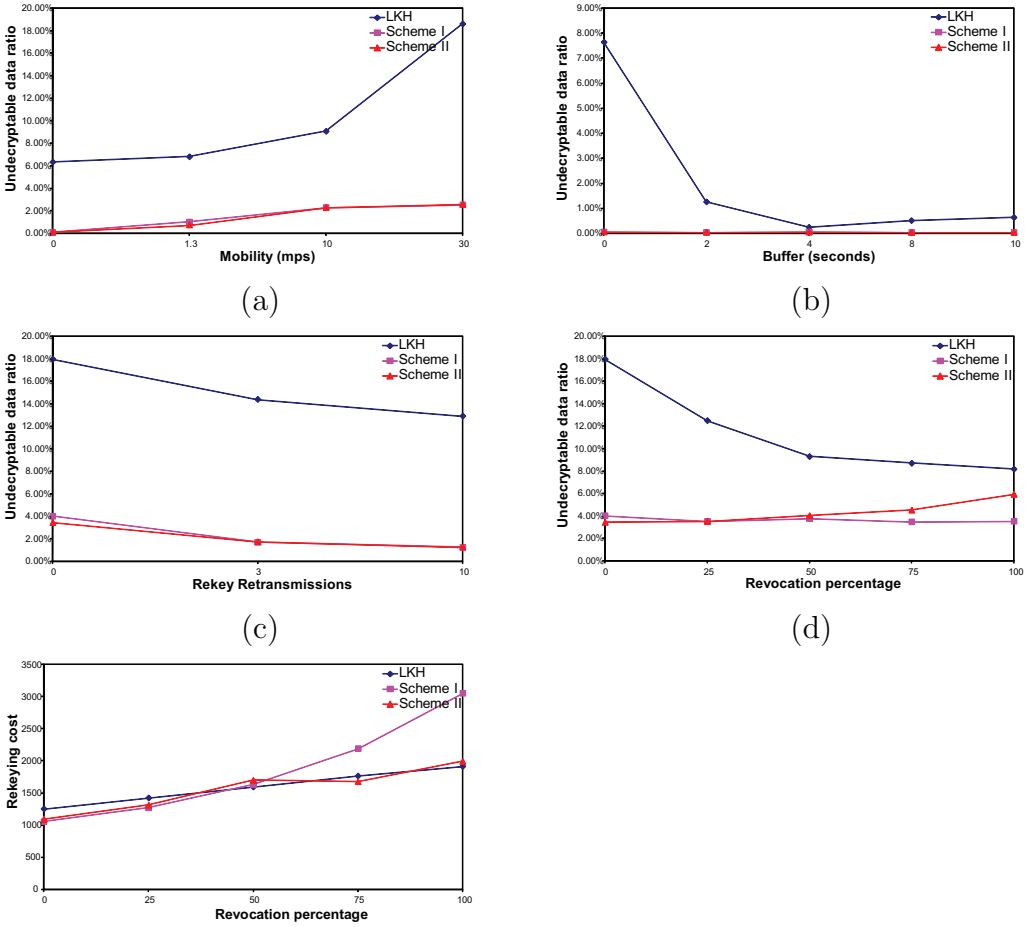


Figure 5.7 Performance evaluation results

the mobility, the higher the number of missed rekey messages. Because LKH is stateful while the proposed schemes are m-stateless, as long as the number of consecutive missed rekey messages is below or equal to m, the self-healing schemes greatly outperforms LKH.

The buffer enables data messages storage for a specific time that to allow members to decrypt previously unsuccessful decrypted data messages. First, note that it may be irrelevant for applications with strict real-time requirements to decrypt previous data. Second, buffering needs storage capacity dedicated for every multicast stream which may vary with the members' devices or simply be absent. As shown in figure 5.7b, it only really benefits LKH mainly because the self-healing period make members know in advanced m decryption keys, therefore limiting the use of the buffer. The influence of the reliable key delivery methods by statically setting a number of retransmissions for every sent key also share a similar behavior as suggests figure 5.7c.

Finally, the impact of the percentage of members who get revoked on the undecryptable data ratio and the communication overhead is studied. LKH increase in performance seems with a stronger number of member revocation can be explained by an increase of rekey messages when more of them are issued by the GCKS. Figure 5.7d also shows that scheme I, characterized by a continuous self-healing property offers a more constant ratio than scheme II which resets the self-healing period after every batched revocation process. Note that the overhead includes the number of keys sent by multicast and unicast transmissions triggered either by an event or a member's request for keys. As the analytical analysis suggests, the communication overhead of scheme I increases exponentially when the keys owned by logical nodes are flagged which is caused by revocation of members. The results shown in figure 5.7e confirm the pattern.

The empirical results show a major improvement over the stateful LKH in scenarios where packet loss is non-negligible such as in typical mobile scenarios. The proposed solution addresses urgent and critical needs in providing secured multicast flows for users using mobile devices with low resources in experiencing fluid multimedia flows even while managing handovers. The two schemes also provides flexbility for the network operators to meet their security policies.

5.6 Conclusion

In this paper, two self-healing schemes based on the DDHC have been proposed for LKH to tackle secure multicast in a mobile environment. Although its weakness in collusion attacks and overhead increase after revocation events, scheme I can be suitable for real time applications with loose security requirements where passed data have little interest and the

number of revocation is low. Such services may include the stock market quote distribution or weather forecast. A more robust and scalable collusion-free approach has been proposed in scheme II in which the self-healing property is maintained between two revocation events. To maximize the self-healing period, batched revocations with variable timeouts have been used. Results show major improvements over LKH in the ratio of undecrypted data messages and bandwidth overhead.

CHAPITRE 6

SECURE ROUTE OPTIMIZATION FOR MIPV6 USING ENHANCED CGA AND DNSSEC

Angelo Rossi and Samuel Pierre
angelo.rossi@polymtl.ca samuel.pierre@polymtl.ca
Mobile Computing and Networking Laboratory (LARIM)
Ecole Polytechnique de Montreal
Montreal, H3T 1J4 Canada

Suresh Krishnan
suresh.krishnan@ericsson.com
Ericsson Research
Town of Mount Royal, QC H4P 2N2, Canada

Abstract

At the moment, nearly half of all Internet subscribers come from mobile units and it is expected to be the largest pool of Internet users by the next decade. The most obvious choice for mobile operators to support more users would be to replace Mobile IP for IPv4 with MIPv6. However, much work is required for its most attractive feature, the Route Optimization (RO) mechanism, to be efficient and secure. RO replaces the inefficient triangle routing by allowing a mobile node (MN) to bidirectionnally communicate with the corresponding node (CN) without passing through its home agent (HA). The lack of pre-shared information between the MN and the CN makes security in RO a difficult challenge. MIPv6 adopts the return routability (RR) mechanism which is more to verify the MN reachability in both its home address (HoA) and care-of address (CoA) than a security feature. Other works attempted to solve the multiple security issues in RR but either their design are flawed, or rely on unrealistic assumptions. This work first presents an enhanced cryptographically generated address (ECGA) for MIPv6 that integrates a built-in backward key chain and offers support to bind multiple logically-linked CGAs together. ECGA tackles the time-memory tradeoff attacks while being very efficient. It is part of the proposed secure MIPv6 (SMIPv6) with secure and efficient RO which uses DNSSEC to validate CGAs from trusted domains and provide strong authentication rather than sender invariance. The AVISPA on-the-fly model

checker (OFMC) tool has been used to show that the proposed solution has no security flaws while still being lightweight in signalling messages in the radio network.

Keywords: Route Optimization, Return routability, cryptographically generated address, Mobile IPv6, DNSSEC

6.1 Introduction

The fight between MIPv6Johnson *et al.* (2004) and PMIPv6Gundavelli *et al.* (2008) as the chosen standard in the mobile industry to replace MIPv4Perkins (2002) has never been fiercer. The differences between those two mobility protocols are more than just technical, their essence on which they were conceived are completely different. On one hand, the user in MIPv6 has complete control of its operations because the network is completely passive and only the MN can initiate any action. On the other, PMIPv6, with the introduction of the mobile access gateway and the local mobility anchor, implements many mobility functionalities through the network. Because the focus in this paper is on MIPv6, it is important that the proposed Route Optimization (RO) mechanism fits in the MN-initiated actions' philosophy.

Two modes are available for data to be transferred between the MN and the CN. First, the bidirectional tunnelling (BT) (Figure 6.1) sends data in an IP-in-IP tunnel using the MN's HA as the intermediary entity that encapsulates and decapsulates the received data. Although inefficient especially when the MN and the CN are relatively near each other compared to the HA, BT is the preferred option for the operators because it does provide more control or monitoring for billing or other purposes.

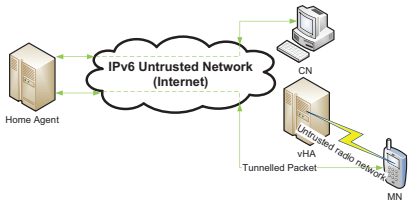


Figure 6.1 Bidirectional tunneling

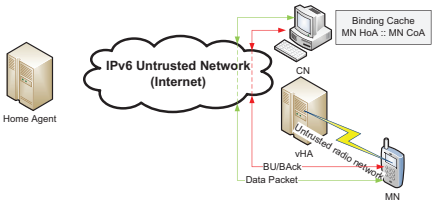


Figure 6.2 Route Optimization

RO (Figure 6.2) enables the data to be exchanged directly between the MN and the CN where the MN fills the MIP destination option with its HA and the CN uses the routing type 2 to send packets to the MN. Prior in exchanging data directly between the MN and the CN, the return routability (RR) Johnson *et al.* (2004) (Figure 6.3) mechanism is executed to test the reachability of the MN's HoA and CoA. More precisely, the MN initiates the RR by sending the Home Token Init (HoTI) tunnelled through its HA and the Care-of Token Init (CoTI) directly to the CN, both carrying a distinct init cookie which is later returned.

Upon reception, the CN generates the HoT and CoT each containing a cleartext keygen token generated by taking the first 64 leftmost bits of H(K_{CN} |HoA|N_i |0) and H(K_{CN} |CoA|N_j |1) respectively where N_i and N_j are the nonces and K_{CN} is a 20 bytes long random string kept secretly by the CN. Once the MN receives the HoT and CoT messages from the CN, it hashes the concatenation of both keygen tokens to generate the binding management key (K_{bm}) which is used to sign binding updates (BU) to the CN. It is important to note that, as stated previously, RR is not meant to secure the RO process, but mainly to test the reachability of both addresses of the MN. Consequently, multiple vulnerabilities Nikander *et al.* (2005); Kavitha *et al.* (2010) can be conducted to the MIPv6 RO mechanism or its optimized version W. Haddad (2005).

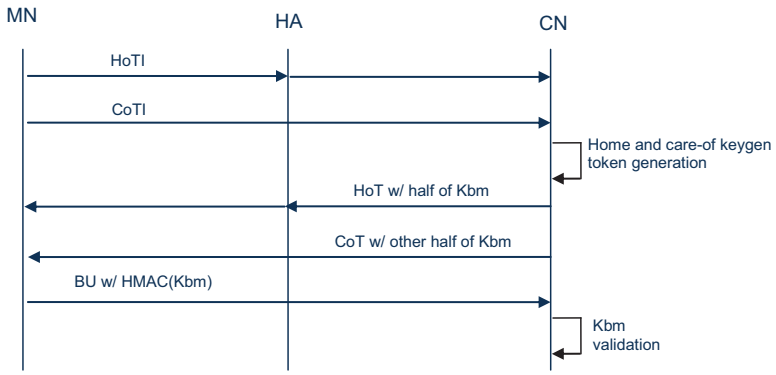


Figure 6.3 Return routability in MIPv6

This paper presents secure MIPv6 (SMIPv6), a secure and efficient RO mechanism for MIPv6 based on cryptographically generated address (CGA) and domain name service security extensions (DNSSEC) which has been verified through an on-the-fly model checker (OFMC). Section 6.2 first points out the security issues with RR and then discusses and analyzes the available alternatives. Section 6.3 and 6.4 detail the Enhanced CGA (ECGA) and the secure RO for MIPv6 using ECGA and DNSSEC. The AVISPA guidelines of protocol implementation along with the results then follow in Section 6.5. Section 6.6 presents the security analysis of the proposed solution and Section 6.7 concludes with a brief summary outlining key observations.

6.2 Background concepts and related work

6.2.1 Security issues with RR

The main scope of RR is to ensure that the MN is reachable in the stated HoA and CoA and, only as a secondary objective, it provides a rudimentary way to authenticate and option-

ally provide confidentiality assuming the symmetric key sent in clear text through the HoT and the CoT has not been intercepted. Obviously, such assumption makes RR unsuitable to a typical network architecture in which messages are exchanged through untrusted networks such as the internet or any other shared network. The following presents the most important security threats in RR.

Session hijacking and Man-In-The-Middle (MITM) attacks

It is sufficient to sniff only the HoT from a CN for an attacker to forge and send a CoTI with its own IP instead of the MN's CoA. The CN therefore responds with the CoT which the attacker uses in combination with the intercepted HoT to generate a valid binding update management key (K_{BU}). It can then send a legitimate BU to the CN to redirect the flow towards him and act as an intermediary between the MN and the CN and read all the messages.

Denial of service (DoS) and flooding attacks

DoS attacks typically exploit limitation in a target's resources and therefore, mobile nodes, thus the MNs and possibly even the CN, are the most vulnerable because of their limited computation power, memory space and bandwidth. An attacker could dramatically increase the CN's workload by intercepting N_H home and N_C care-of tokens and send $N_H * N_C$ legitimate BUs forcing the CN to store many sessions in memory and redirect multiple flows in a very short period of time. Note however, that in order to make this attack successful, the forged BUs must be sent to the CN before the nonces are refreshed. Network flooding DoS attacks can also be led on MNs or, more broadly, on a visited network if an attacker manages to redirect only signaling MIPv6 messages without being in the data path. Such a favorable environment enables the attacker to send forged HoTIs and CoTIs with any HoA and CoA and intercept the related keygens to redirect the flows where it wishes. A similar but easier alternative would be to enter a visited network, establish multiple RO sessions and simply detach from its care-of link while sending acknowledgements to the CNs to make them believe that it is still active on its CoA W. Haddad (2009).

6.2.2 Certificate-based RO protocols (CBU and HCBU)

Authors in Deng *et al.* (2002) propose two RR alternatives by introducing a certificate scheme and only make the trusted entities execute RO related operations. First, the Certificate-based Binding Update (CBU) protocol makes use of a Certificate Authority (CA) that issues a certificate for every home link subnet prefix (HLSP) where the private key is

kept secretly by the HA. As depicted in Figure 6.4, upon receiving a RO request and after exchanging nonces with the CN, the HA provides its HLSP and public key through its certificate signed by a trusted CA. It then starts an authenticated Diffie-Hellman (DH) exchange by signing EXCH0 with its private key in order to share a secret key K_{DH} which is then used to generate K_{BU}.

CBU's design has many flaws starting with the disclosure of the secret value x used in the computation of K_{DH} that an attacker can easily compute by applying a logarithmic function ($x=log_g(g^x)$) if the public components g, g^x are intercepted. Also, the lack of validation of the CoA enables a legitimate but malicious MN to spoof any victim's IP or subnet leading to redirection attacks. Secondly, the assumptions are unrealistic and impractical. The presence of fragmented authentication infrastructures across different domains is necessary to let the CN validate the CA that issued the HLSP certificate, and thus avoid MITM and impersonation attacks. Yet, this poses scalability and flexibility issues in trust management when expanding this flat structure into a more global approach. Furthermore, the different administrative domains are typically reluctant in sharing information or accept trusted CA certificates from competitors.

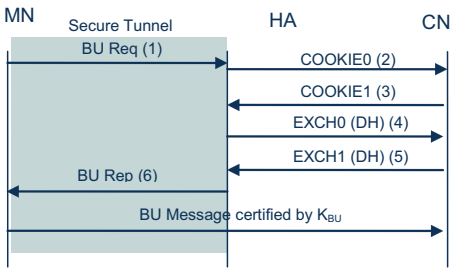


Figure 6.4 Route Optimization in CBU

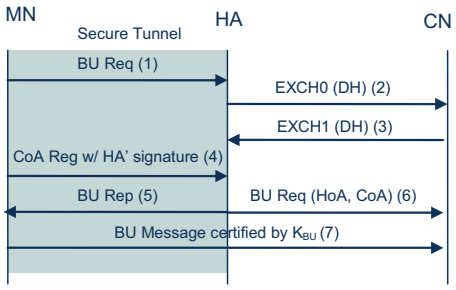


Figure 6.5 Route Optimization in HCBU

To tackle the CBU issues, the same authors proposed the Hierarchical Certification-based Binding Update (HCBU) that uses a 3-layer chain certification scheme in which the root CAs (known by everyone) (layer 1), the intermediate ISPs (layer 2) and finally the MN's domain (layer 3) sign the 64-bit long IPv6 HoA and CoA subnet prefix of the MN. The trust delegation enables the CAs in the operator's domain to generate chain certificates for the MN's HoA and CoA which can be validated up to the global trusted CAs. Also, the MN must supply a proof of ownership of its CoA signed by the visited HA (vHA) to its HA prior the establishment of RO (Figure 6.5).

Although HCBU (or its optimized version D. Kavitha (2010)) is certainly more secure than CBU, the authors' assumptions are unrealistic. First, the global deployment of a 3-layer chain certification of IPv6 subnet prefix is currently nonexistent and such implementation requires

much structural changes and collaboration between many consortiums. Because there is no other use for it, the incentive for such approach is very limited. Second, HCBU does not verify if a bidirectional route exists from the CN to the MN prior to the acceptation of the BU and redirection of the flow. In fact, links across the internet core are composed of numerous automated systems having different policies (QoS, security, etc.) that may force upstream and downstream flows to travel in different paths between the same 2 nodes, or worse, deny one of them. Third, the authors assume that the security policies between the vHA and HA or between the CN and the HA are always compatible, thus avoiding the scenarios where the HA generates a key too long for the CN or the vHA signature is non compliant with the HA minimal requirements. Beside the assumptions, HCBU is still vulnerable to the network flooding attack described in W. Haddad (2009).

6.2.3 Cryptographically Generated Address (CGA)

The scope of a CGA is to prevent spoofing attacks by binding the suffix of an IPv6 address with a self-generated certificate and signing the hashed message authentication code (HMAC) of each message with its private key. More precisely, CGA provides sender invariance Drielsma *et al.* (2007) in which participants that initially share no relationships at all, to authenticate messages only once the link between the CGA and its certificate has been acknowledged by the other parties. In fact, any node, including an intruder, can generate a pseudonym (in this case the CGA), which may not be linked to the real identity of the node. This may lead to a MITM attack if the initial messages between two nodes are tempered with such in S. Bradner (2003). Note, however that an intruder is unsuccessful in signing or decrypting messages belonging to someone else's pseudonym. Sender invariance is therefore weaker than authentication where the real identities between participants are known prior communicating and do not rely on an initial leap of faith about the pseudonymous identity.

CGA Generation

After generating its public/private key pair, the node constructs a random 128-bit modifier to compute hash-1 using a given H hash function:

Hash-1 = H(Modifier | Subnet | Collision count | Public Key)

The first 64 bits of a CGA IPv6 address are reserved for the subnet while the remaining 64 bits are the first leftmost 64 bits of hash-1. By providing its CGA parameters (components in hash-1), a node's CGA can be validated with a single hash operation and the message can be validated through the HMAC signature. Having generated its own certificate, an attacker can perform a bruteforce attack on its modifier to match a victim's CGA and steal

its pseudonym. CGA must therefore be resilient to such attack as it will break the sender invariance property. Because 5 bits of a CGA are reserved, an attacker requires a maximum of 2^{59} hash operations to spoof a CGA given a known modifier and certificate. To increase the strength against such attack, the hash extension technique has been proposed where hash-2 is computed prior to hash-1. The scope is to increment the modifier until the results is 16*SEC bits of 0:

Hash-2 = H(Modifier | 9 bytes of 0 | Public Key) = 16$*$SEC bits of 0

The value of the SEC parameter varies between 0 and 7 and is directly related to the strength of the CGA. As shown in Table 6.1, the higher the SEC value is, the higher the computational cost to generate the CGA. Consequently, the number of required hash operations is increased by a factor of 2^{16*SEC} which brings the maximum operations for a successful brute-force attack from 2^{59} to $2^{59+16*SEC}$.

Table 6.1 Computation time of CGA on a AMD64 processor according to Bos *et al.* (2009)

SEC value	1	2	3
Required Time	0.2 sec	3.2 hrs	24 yrs

Time-memory tradeoff attack

Because hash-2 is computed with invariant elements that are not related to the node's network attachment, suitable modifiers for a given SEC value can be pre-computed over time. In fact, assuming no hash-1 collision, a table listing at least 2^{59} pre-computed valid modifiers for known public/private key pairs would guarantee an attacker to hit a victim's CGA with 2^{59} hash operations independently of the SEC value, disabling the hash extension security gain. An easy solution would be to include the subnet in both hashes (hash-1 and hash-2), thus greatly complicating the attack by forcing a pre-computed table for each available subnet. However, such approach forces a node to recompute its hash-2 every time it changes subnet. As shown in Table 6.1, for SEC=2, it will require a node to wait several hours before it can attach to the network. In a mobile environment where subnet migration is frequent, this delay is not tolerable.

Replay attack

The authentication in CGA relies in the digital signature added at the end of each sent messages, but the fact that the IP header (the source address) is excluded from the signature

opens the door to several replay attacks as detailed in Bos *et al.* (2009). The easiest one is for an attacker to sniff and store messages of its victim while obtaining its modifier and public key. The attacker then changes subnet, regenerates the victim's CGA by recomputing hash-1 with the victim's CGA parameters and replay the stored messages after changing the source address.

6.2.4 CGA++

In Bos *et al.* (2009), authors propose small modifications to CGA in order to tackle the time-memory tradeoff attack issue by providing authentication in the computation of the CGA. To do so, they first include the subnet in the computation of both hashes. After hash-2 is computed, a signature of the modifier, collision count and subnet prefix is concatenated with the corresponding public key and then hashed to form hash-1. As discussed earlier, including the subnet in hash-2 requires the MN to regenerate its CGA every time it changes subnet which imposes prohibitive delays in a mobile environment. The authors chose to do so to counter the time-memory tradeoff attack which in theory would require 2^{64} tables to cover all the available subnets in IPv6. However, in practice, because subnets are statically configured and remain unchanged once they have been assigned to the operator by IANA, there is a higher risk that subnets used in major network operators who have a large proportion of customers are targeted first by attackers. By focusing on the most popular subnets, the number of required pre-computed tables is greatly reduced while the number of potential victims remains large.

6.2.5 Enhanced Route optimization for MIPv6 (RFC 4866) and other CGA-based RO protocols

As an alternative to certificate-based RO, authors in Arkko *et al.* (2007) combined the cryptographically generated home address (CGHoA) introduced in M. Roe (2002); W. Haddad (2005) with the existing RR concepts. They also propose more optimized ways (Figure 6.6) to reduce latency by introducing semi-permanent security associations (SA) and credit-based authorization to enable early BUs. The former mechanism establishes a renewable permanent shared secret token (a symmetric key known as the permanent home keygen token) between the MN and the CN to avoid using resources validating the computationally expensive CGHoA (asymmetric) for future BUs. Note however that the first BU to a CN must include the CGHoA's signature to prevent an attacker that intercepted a HoT to forge a BU with it.

The credit-based authorization allows the CN to send packets to a non-tested MN's CoA

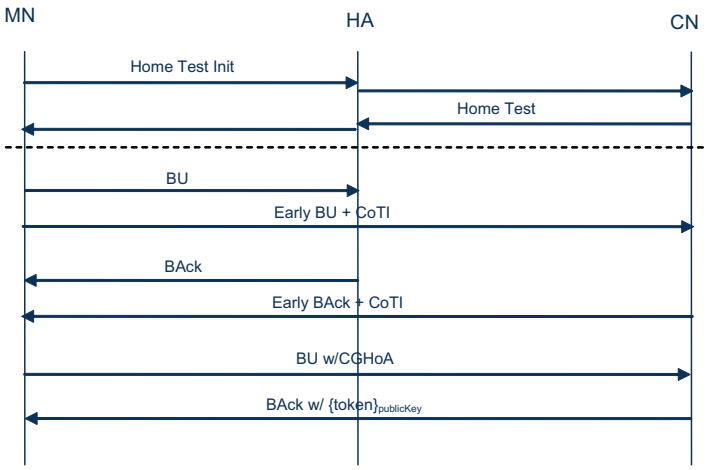


Figure 6.6 Correspondent node registration with authentication based on reachability verification at the home address with concurrent care-of address test

(unverified state) until no more credits for the MN is available. Credits are added when payload packets from the MN's CoA are received by the CN while being in VERIFIED state. This enables the MN to trigger the early BU technique (after a handover) in which the flow is redirected to the new care-of address, while the mobile node's reachability at the new care-of address is verified concurrently. The CN moves the care-of address to VERIFIED state once reachability verification completes.

This solution adds the sender invariance property to the BU in RR. However, an intruder can achieve a MITM attack by intercepting the initial signalling messages from a MN, exchanging with the MN's destination using the attackers own generated CGHoA and responding back to the MN. Because the MN ignores the true identity of the CN, it will validate the attacker's responses.

Also, it offers no way for the CN to verify the ownership of the MN's CoA other than using the CoT which main purpose is to test the CoA reachability. Therefore, a malicious but legitimate MN could redirect flows to any IP by spoofing the CoA. The credit-based authorization mitigates the effects, but can be problematic when a legitimate MN is unable to perform the reachability test in time because of a temporary problem (handover delay, congestion, interference, etc.) and sees its flow cut off. Also as with any reputation/credit based system, the spread between the thresholds can always be exploited by an attacker who well behaves to gain credits and then spends them on flooding a victim. By coordinating this attack with multiple other malicious and colluding nodes or zombies with valid CGAs, a distributed denial of service (DDoS) could easily take a target down. Finally, operators might be reluctant in using a protocol with such a high signalling overhead in the radio access

network.

6.2.6 Other solution-related protocols

To conclude this section, DNS security extensions (DNSSEC) Arends *et al.* (2005a,c,b) and the flush request signalling W. Haddad (2009) are briefly introduced. Authors in W. Haddad (2009) propose a simple solution to repel network flooding aimed at a malicious MN who, after initiating an exchange with the CN, detaches itself from the visited network but keeps regularly sending ACK to the CN to keep alive the flow. To achieve this, the authors first assume that a "symbiotic" relationship exists between the MN and the network it resides in by means of CGA and secure network discovery protocol (SeND)Kempf et Koodli (2008). When the AR detects that the MN is not present anymore in its network, it may immediately, or after a timeout, send a signed Flush Request (FR) with a proof of relationship to the CN that keeps sending messages to the disconnected MN. Note that W. Haddad (2009) presents a concept more than a complete solution.

DNSSEC proposes changes to the DNS protocol to tackle the cache poisoning attack where a malicious node responds to a DNS query before the DNS server claiming that the requested full qualified domain name (FQDN) points to a forged IP address. DNSSEC enables authentication of DNS responses through digital signatures and public-key cryptography. A DNS server authoritative of a domain is trusted through a chain of trust leading to a set of known and trusted DNS root zone. The deployment of DNSSEC at the DNS root was implemented on July 15^{th} 2010 and while the .org top-level domain already offers DNSSEC, a full deployment is expected to be completed by the end of 2011.

6.3 Enhanced CGA (ECGA)

This section presents ECGA which is part of the proposed secure RO in SMIPv6. It details the generation of an ECGA and the support for binding multiple CGAs together.

6.3.1 Notation for ECGA

$\{M\}_K$: Message M is encrypted with key K
X | Y: Concatenation of components X & Y
$H^x(M)$: The x^{th} (default is 1) hash of message M
PK_A: Private key of entity A (asymmetric)

6.3.2 The construction of Enhanced CGA (ECGA): Hash-2 and Hash-1 computation

A backward key chain of length $p \in N$, as shown in Figure 6.7, is constructed by applying a hash function H repeatedly on a secret key s:

$H^p(s), H^{p-1}(s), ..., b_i, ..., H(s)$where $b_i = H(b_{i-1}) = H^i(s)$ and $1 \leq i \leq p \ \forall i, p \in N$

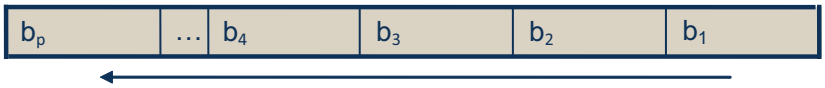


Figure 6.7 Example of a backward hash chain

It is important to note that while the backward key chain is constructed from right to left, the keys must be used from left to right to provide authentication. In fact, any node intercepting b_i can generate b_j for $i \leq j \leq p$ (the left part of the chain) by applying the hash function repeatedly, but, it is computationally infeasible to find the right part b_j for $1 \leq j \leq i$ where $i \leq p \ \forall i, j, p \in N$. For example, let TW be the time window in which a key must reside in, a party that previously received a backward key b_i authenticates the key b_j, for $j < i$, if $H^x(b_j) = b_i$ where $x \leq TW \ \forall x, i, j \in N$. No other node could have found b_i prior to its disclosure because $H^{-x}(b_j)$ is very complex to compute. In other words, for a node that previously received the backward key b_i and wants to validate the current backward key b_j, it simply hashes b_i until finding a match with b_j. If not match has been found within TW maximum allowed hashing operations, the validation process fails.

Once the backward key chain is constructed, both hashes in CGA use current backward key b_i and are generated the following way:

Hash-2: H(Modifier | b_i | Public Key) = 16∗SEC bits of 0

Hash-1: H(Modifier | b_i | Subnet | Collision count | Public key)

As soon as a valid modifier in hash-2 (using b_i) has been found, the node may start to compute the next modifier in hash-2 using the following backward key b_j from the chain, where $j < i$, $\forall i, j \in N$. Once the new modifier has been found for b_j, both parameters are used in the next CGA to generate and b_i is removed from the chain if it is not used by any of its addresses. The operation is then repeated with the next backward key, and so on. This removes the need for a node to know the subnet it is going to attach to before starting to execute the hash extension technique.

6.3.3 Binding multiple CGAs together

ECGA offers support to bind multiple logically-linked CGAs together through cascading them or by simply adding them independently in the hash-1 computation. This is especially

useful to bind the MN's HoA and CoA addresses. Table 6.2 presents an example of CGA generation using the cascading mechanism:

Table 6.2 Example of cascading CGAs

CGA	Hash-1
CGA_1	H(Modifier$_1$ | b_{i1} | Subnet | Collision count | PublicKey$_1$)
CGA_2	H(Modifier$_2$ | b_{i2} | Subnet | Collision count | CGA_1 | CGA_1 Parameters | PublicKey$_2$)
CGA_3	H(Modifier$_3$ | b_{i3} | Subnet | Collision count | CGA_2 | CGA_2 Parameters | PublicKey$_3$)

In the previous example, CGA_3 integrates CGA_2 which itself integrates CGA_1 into its hash-1 computation, binding all 3 CGAs. For a signature to be valid for CGA_3, it must contain a proof of all 3 entities such as a HMAC signed by all 3 parties in the reverse order: $\{\{\{\text{HMAC}\}_{PK3}\}_{PK2}\}_{PK1}$. However, it may be that node 3 is not known to node 1 and thus cannot authenticate it. In that case, the HMAC can be signed in the forward order, but because any node could intercept and sign $\{\{\text{HMAC}\}_{PK1}\}_{PK2}$ to pretend that nodes 1 & 2 approved CGA_3, HMAC must contain information specific about node 3.

When node 3 is bound to nodes 1 & 2 but nodes 1 & 2 are not bound together, the inclusion method can be used by adding the CGA of nodes 1 & 2 in the computation of hash-1 for CGA of node 3:
H([modifier$_3$ | subnet | Collision count | CGA_1 | CGA_1 parameters | CGA_2 CGA_2 parameters | publicKey$_3$])
Again, the signature must contain a proof of all 3 nodes, but because nodes 1 & 2 are not logically bound, node 3 may provide 2 signatures, one showing the relation between nodes 1 & 3, and the other between nodes 2 & 3: $\{\{\text{HMAC}\}_{PK3}\}_{PK1}$ and $\{\{\text{HMAC}\}_{PK3}\}_{PK2}$.

6.4 Secure route optimization in SMIPv6 using ECGA and DNSSEC

The main idea behind using DNSSEC and ECGA is to authenticate a message according to the domain it originated from. This section first presents the objectives and assumptions of the proposals. Second, the integration of ECGA in the SMIPv6 is detailed. Finally, the notation used in describing the secure RO proposal fllows along with the detailed execution in various MIPv6 scenarios leading to the RO.

6.4.1 Objectives

The objectives of the secure RO proposal are :

- Be as secure as HCBU while relying on realistic assumptions;
- Lean towards providing authentication globally rather than the sender invariance by validating the CGA (pseudonym) through the use of DNSSEC and trusted domains;
- Provide more flexibility and control to mobile operators when allowing RO.

6.4.2 Assumptions

For the proposed solution to strongly authenticate the MN's HoA and CoA to the vHA and the HA respectively, the following assumptions must be met:

1. 64 bits are sufficient for the care-of subnet;
2. Whitelist of trusted domains are kept up to date by the operator administrators (in HA and vHA);
3. DNSSEC is deployed globally;
4. Operators protect their network by analyzing and filtering all incoming and outgoing messages.

To provide global strong authentication, an additional assumption is necessary

5. A trusted authority that publishes a list of trusted mobile operator domains

6.4.3 ECGA in SMIPv6

When a MN first boots, it starts by constructing a backward key chain long enough to support many creation of CGCoAs and avoiding refreshing its CGHoA too often. Table 6.3 shows how CGAs are generated in SMIPv6 to provide secure RO while Table 6.4 presents their related signature. Note that b_i and b_j are the backward keys of the MN from the same chain where $1 \leq j \leq i \ \forall i, j \in N$.

6.4.4 Notation for the secure RO solution

On top of the notation used for ECGA 6.3.1, these are used for the RO solution:

PSK_{A-B}:Preshared secret key between A & B
K_{A-B}: Symmetric private key between A & B
M(X | Y)Message M sent with parameters X & Y

Table 6.3 ECGA in SMIPv6

CGA	Hash-1
CGHA	H(Modifier$_{HA}$ | Subnet | Collision count | PublicKey$_{HA}$)
CGVHA	H(Modifier$_{VHA}$ | Subnet | Collision count | PublicKey$_{VHA}$)
CGHoA	H(Modifier$_{HoA}$ | b_i | Subnet | Collision count | CGHA | CGHA parameters | PublicKey$_{MN}$)
CGCoA	H(Modifier$_{CoA}$ | b_j | Subnet | Collision count | CGVHA | CGVHA parameters | CGHoA | CGHoA parameters | PublicKey$_{MN}$)

Table 6.4 Signatures used for secure RO in SMIPv6

CGA	Signature
Home Token	$\{$H(CGHoA | CGHoA parameters)$\}_{PKHA}$
Visited Token	$\{$H(CGCoA | CGCoA parameters)$\}_{PKVHA}$
Auth Token	$\{\{\{$H(CN | CGCoA | CGCoA parameters)$\}_{PKMN}\}_{PKHA}\}_{PKVHA}$

6.4.5 Bootstrapping in home network

When the MN first bootstraps into its home network, it must approve its generated CGHoA to its HA and in return, the HA registers a MN's FQDN linked to its CGHoA using DNSSEC and sends the MN the home token used when attaching to a visited network. Figure 6.8 details the message sequence.

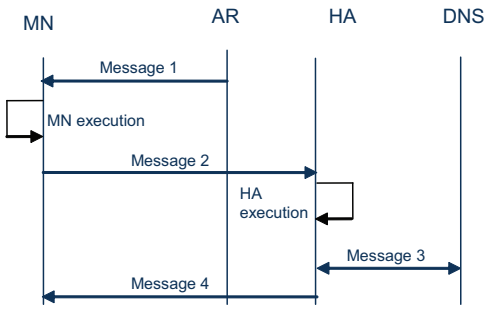


Figure 6.8 Bootstrapping process of a MN in its Home Network

6.4.6 MN's attachment to visiting network

When the MN enters a visiting network (Figure 6.9), the execution is similar to its home network's bootstrap with the exception that the signature request is authenticated through

Message 1:	AR → MN : Router Advertisement(Home subnet | CGHA) | CGHA Parameters | $\{HMAC\}_{PKHA}$
MN Execution:	MN generates its CGHoA using its CGHA and subnet
Message 2:	MN → HA: Signature Request(CGHoA) | CGHoA Parameters | $\{HMAC\}_{PSKMN-HA}$ | $\{HMAC\}_{PKMN}$
HA Execution:	HA authenticates the message through CGHoA validation and preshared info and generates the signed home token
Message 3:	HA ↔ DNS Server : using DNSSEC, the HA registers the MN's HoA FQDN and links it with CGHoA
Message 4:	HA → MN: Signature Response(home token| HoA FQDN) | CGHA Parameters | $\{HMAC\}_{PKHA}$

the signed home token and the MN's public key used for its CGHoA.

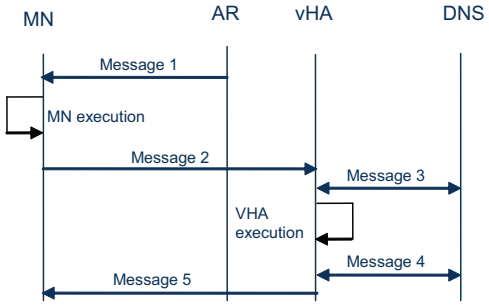


Figure 6.9 MN enters a new visiting network

Message 1:	AR → MN: Router Advertisement(Visited subnet|CGVHA)|CGVHA Parameters|$\{HMAC\}_{PKVHA}$
MN Execution:	MN generates its CGCoA using the claimed CGVHA and subnet
Message 2:	MN → VHA: Signature Request(CGCoA| home token)|CGCoA Parameters|$\{HMAC\}_{PKMN}$
Message 3:	VHA ↔ DNS Server: using DNSSEC, the VHA executes a forward-confirmed reverse DNS (FCrDNS) on the CGHoA and CGHA and makes sure their domain is the same
VHA Execution:	VHA first that the MN home network's domain is trusted and then authenticates the message through the validation of CGHA, CGHoA and home token, and if successful, generates the signed visited token
Message 4:	VHA ↔ DNS Server: using DNSSEC, the VHA registers the MN's CoA FQDN and links it with CGCoA
Message 5:	VHA → MN: Signature Response(visited token |CoA FQDN)|CGVHA Parameters|$\{HMAC\}_{PKVHA}$

Once the MN successfully attaches to the visited network, it must send a BU to its HA and alert it of its new CGCoA (Figure 6.10).

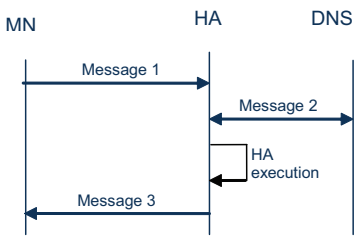


Figure 6.10 Binding Update to HA

Message 1:	MN → HA: BU| visited token |CGCoA Parameters|{HMAC}$_{PKMN}$
Message 2:	HA ↔ DNS Server:using DNSSEC, the HA executes a FCrDNS on the CGCoA and CGVHA and makes sure their domain is the same
HA Execution:	the HA first checks its whitelist of allowed domains and then authenticates the message through the validation of CGVHA and CGCoA by making also sure that the backward key used for CGCoA is part of the chain. The HA could also force its MN to change to the next backward key if the time allowed to use that key exceeds the one specified in its security policies
Message 3:	HA → MN: BAck|CGHA Parameters|{HMAC}$_{PKHA}$

6.4.7 Secure RO for SMIPv6

For the CN to validate the CGCoA, it requires a signature involving the MN, the VHA and the HA. Figure 6.11 details the message sequence of the proposed solution where Auth Token=H(CN|CGCoA|CGCoA Parameters)

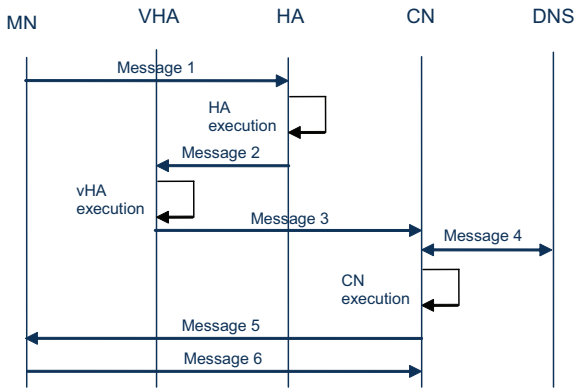


Figure 6.11 Secure and efficient route optimization message sequence

The symmetric key K$_{MN-CN}$ along with the signed visited token are used to authenticate the MN's new CGCoA to the CN when the handover has been executed inside the same visited domain. Otherwise, the Auth Token must be signed by the new vHA and thus becomes invalid and the RO execution must start from the beginning.

Message 1:	MN → HA: RO Authentication(CN|CGCoA|CGCoA Parameters|{Auth Token}$_{PKMN}$)|CGCoA Parameters|{HMAC}$_{PKMN}$
HA Execution:	Validation of the Auth Token and its policies to see if RO is allowed for the MN and the visited domain
Message 2:	HA → VHA: RO Authentication(CN|CGCoA|CGCoA Parameters|{{Auth Token}$_{PKMN}$}$_{PKHA}$)|MN credentials|CGCoA Parameters|{HMAC}$_{PKHA}$
VHA Execution:	Validation of the Auth Token and if credentials received comply with its policies, sign the token and send it to the CN
Message 3:	VHA → CN: RO Authentication(CN|CGCoA|CGCoA Parameters|{{{Auth Token}$_{PKMN}$}$_{PKHA}$}$_{PKVHA}$)|CGVHA Parameters|{HMAC}$_{PKVHA}$
Message 4:	CN ↔ DNS Servers: using DNSSEC, the CN executes a FCrDNS to ensure that the CGHoA and CGHA are on the same trusted domain. It repeats the operation for the CGCoA and the CGVHA
CN Execution:	The CN validates all CGA addresses while making sure the CGHoA and its parameters are the same then the current one used for bidirectional tunnelling (BT). Then, it generates a symmetric key K$_{MN-CN}$
Message 5:	CN → MN: {AUTH Response(K$_{MN-CN}$)}$_{PublicKeyMN}$
Message 6:	MN → CN: BU|{HMAC}$_{KMN-CN}$

6.4.8 Flush Request for SMIPv6

When the MN detaches from its visited network while leaving ongoing flows active, the VHA informs the sources to stop sending their flow. The Flush Request (FR) message is sent from the VHA to the CNs:
FR(CN | CGCoA | CGCoA Param)| CGvHA Param | CGvHASign

6.5 AVISPA implementation guidelines and results

The bootstrapping executions in the home and visited networks, the BT and the RO of SMIPv6 have been implemented in HLPSL. This section briefly presents the key ideas on how the proposed solution was implemented in AVISPA Armando *et al.* (2005) model checker. Secondly, the results show that the solution respects the stated security goals.

6.5.1 Security Goals

The proposed solution implements all security goals available in AVISPA:

1. weak authentication to ensure sender invariance of the router advertisement messages;
2. strong authentication for all other messages sent from the MN, HA, VHA and the DNS server;

3. secrecy to ensure confidentiality of symmetric keys held by the MN and shared with its HA and the CNs.

DNSSEC enables a node to securely verify if a CGA is part of a trusted domain. Given that a CGA cannot be spoofed, receiving a message from a CGA signed by a trusted entity (HA/VHA) of a trusted domain increases the security property for messages from sender invariance to authentication. However, because no assumption has been given about the knowledge of trusted domains to the MN, it must rely initially on the sender invariance property for the router advertisements.Note that the sender invariance is implemented in AVISPA using the *weak_authentication_on* goal along with a *witness* and *wrequest* security predicates using the message's source identity as the authenticator and the authenticated. As an example, the home router advertisement uses the (HA,HA,<goal id>, CGHA's signature) parameters.

6.5.2 Roles and their initial knowledge

In AVISPA, the message sequences along with the security predicates are confined within roles and represent the core of the implementation. Figure 6.12 shows the initial knowledge for each roles in the route optimization of SMIPv6 and how they have been instanciated.

6.5.3 Sessions and intruder knowledge

Figure 6.13 shows the tested scenarios where the intruder took the place of the MN, the HA and the VHA. The knowledge of the intruder includes the PSK with the HA (useful when the attacker acts as a MN), the PSK with the DNS (useful when the attacker acts as a HA), the registered domain owned by the intruder, its registered FQDN, its self-generated certificate and available public information related to the HA, VHA and the DNS server.

6.5.4 AVISPA results

Both back-ends offered in AVISPA, the On-the-Fly Model-Checker (OFMC) and the Constraint-Logic-based Attack Searcher (CL-AtSe), have been used to verify security flaws in SMIPv6 using the Dolve-Yao intruder model. The OFMC is a reactive and demand-based model checker that uses symbolic techniques to convert the states to build an infinite tree out of the analysis of the protocol. On the other hand, the CL-AtSe transforms each step into a set of constraints on the attacker's knowledge which is then tested against the provided set of security goals. As Figure 6.14 shows, both AVISPA back-ends, launched with a bounded number of sessions, conclude that the proposed solution is secure according to the goals in 6.5.1.

6.6 Security Analysis of SMIPv6

As part of the analysis on the proposed solutions, the assumptions are first discussed followed by the compatibility issues of SMIPv6 with MIPv6 extensions, to conclude with the advantages and limitations.

6.6.1 Discussion on the assumptions

64 bits IPv6 subnets

Some works Deng *et al.* (2002); Arkko *et al.* (2007) argue that using CGA for the MN's CoA would interfere with its addressing or that because the MN's CoA is not permanent, there is no interest in being cryptographically generated. While it is true that the MN's CoA change every time it changes subnet, ECGA uses a pre-computed hash-2 and only requires one hash operation to generate its new CGCoA while guaranteeing at least the sender invariance property (to CN) or authentication (to VHA) in untrusted networks. As for interfering with its addressing, the first 64 bits are reserved for the subnet and should typically be enough. If not, because ECGA hash extension is not vulnerable to the time-memory attack, it would be possible to reduce the number of bits dedicated to it and compensating by increasing the SEC.

White list of domains

For an entity to authenticate a message, it is crucial to have a list of domains that it can trust. The mobile operators must maintain a list of domains that allow its MNs to:

- Connect to the domain, update the binding cache and allow BT
- Allow RO

This adds control, flexibility and scalability to operators who can configure its whitelist of domains according to the agreements it holds with its competitors. Wildcards cannot be used as it would let any registered domains to be trusted.

DNSSEC global availability

As opposed to a global scale certificate authorities or subnet certification using chain certificates, DNSSEC is already being deployment today and is scheduled to be available worldwide by the end of 2011.

Mobile operators secure their network

To keep their trusted status, mobile operators have every incentive to secure their network by analyzing every incoming and outgoing messages. If attacks can be lead from an operator, it may loose its trust relationships with other operators and consequently see its customers being deny RO or event BT in visited networks.

Centralized list of trusted mobile operator domains

This assumption is the most questionable as it requires a centralized list of all mobo domains to be published and maintained by a trusted authority. It can be realized in many different ways such as a website or a DNS-based white list DNS (2011).

6.6.2 Compatibility issues with MIPv6 alternatives: FMIP, HMIP, PMIP

An important aspect of SMIPv6 is that the MN-oriented philosophy of MIPv6 in which all operations are initiated by the MN is kept. This allows MIPv6 alternatives like FMIPv6 Kempf et Koodli (2008) and HMIPv6 Soliman *et al.* (2005) to be compatible with the proposed solution. For PMIPv6's Gundavelli *et al.* (2008) network-initiated handovers, the MN's CGCoA must be computed by a network equipment and thus requires a point-to-point radio link to secretly send the private key to the MN or to authenticate it so the network can sign the MN's messages before heading to the untrusted network.

6.6.3 Advantages and limitations

ECGA

To tackle the replay attacks, ECGA includes the MN's CGCoA in the signature, either through the signed visited token or the Auth Token which are required to establish a BT or RO. Therefore, even if an attacker uses the CGCoA parameters in a new subnet, the CGCoA will change and the signatures become invalid. Also, the hash-2 computation in ECGA includes a backward key for the double purpose of weak authentication and prohibiting precomputed hash-2 tables, thus eliminating the time-memory tradeoff attack. Therefore, for an attacker to impersonate another's ECGA, it must not only match the vitim's hash-1 and hash-2, but also find the next backward key that will be used. This additionnal complexity is kept as long as a node changes at least TW times its backward key in the average time required for an attacker to perform a successful impersonation attack. Otherwise, an attacker could use the victim's current (or past if TW>1) backward key and the impersonation complexity falls to the same as CGA.

Moreover, ECGA avoids inflicting additional delays by enabling a node to pre-compute hash-2 before having the subnet it is going to attach to. Therefore, it would be possible to dynamically adjust the SEC parameter according to the rate of generation of new CGAs (which is linked to MN's mobility), the average available processing power and remaining energy. From a security standpoint, a higher value of SEC is always preferred, but the more processing power, energy consumption and computation time is required. Table 6.5 compares CGA, CGA++ and ECGA.

Table 6.5 Comparison of worse case operations in CGA, CGA++ (using a 1024-bit RSA key) and ECGA. All timings are expressed in hash function evaluations. The parameter sec=s is the security parameter used for hash extensions, TW is the validation time window and $|key|$ is the length in bits of a backward key

Metric	CGA	CGA++	ECGA
Time for address generation			
with sec=0	1	$1+2^{10.9}$	1
with sec $\geq$ 1	$1+2^{16*s}$	$1+2^{16*s+10.9}$	$1+2^{16*s}$
Time for address verification			
with sec=0	1	$1+2^{5.5}$	1+TW
with sec $\geq$ 1	2	$2+2^{5.5}$	2+TW
Impersonation time *			
with sec=0	2^{59}	$2^{69.9}$	$2^{59+\|key\|}$
with sec $\geq$ 1	$2^{59+16*s}$	$2^{59+16*s+69.9}$	$2^{59+16*s+\|key\|}$
Address renewal time when moving to a different network			
with sec=0	1	$1+2^{10.9}$	1
with sec $\geq$ 1	1	$1+2^{16*s+10.9}$	1
Resistance against the time-memory tradeoff attack	No	Yes	Yes
Resistance against the replay attack	No	Yes	Yes
Storage requirement for			
validator	None	None	Backward key
generator	None	None	Hash chain
Background address generation	Not required	No	Yes

* Assuming the backward key has been changed at least TW times during the average time it requires for a sucessful impersonation attack

Secure RO proposal for SMIPv6

The authentication in SMIPv6 is based on the domains' ECGA trustiness through DNSSEC and its whitelist. First, the MN must prove ownership of its addresses through ECGA. Sec-

ond, the ECGA must be linked to a trusted entity in a trusted domain and a signed through signed through the home and visited tokens. By receiving such information, a party that trusts both the home and visited domains of the MN, can authenticate the messages as long as no attacks can be lead from within a trusted domain. It is therefore important for a mobile operators to analyze and filter all incoming and outgoing traffic. For example, a malicious MN who spoofs another node's ECGA using its own self-generated certificate to send messages can be easily detected by comparing the public key of the attacker and the legitimate owner.

Because a priori no information about the CN is known, it is only possible to ensure its sender invariance. Therefore, any attacker intercepting the initial request from a legitimate source can act as a CN. Futhermore, if the same attacker comes from a trusted domain, it can authenticate himself to the CN and execute a MITM attack. This security issue is inherent to the anonymous nature of the CN and is impossible to tackle unless pre-shared information is available. Such attack can even be lead by a colluding node in untrusted domains that receives the tokens and certificate of a MN in a trusted domain and that manages to intercept messages going to its partner. On the other hand, as long as a MN from a trusted domain does not share its private key, it would be possible for a trusted domain to block routing type II messages from its wireless network to avoid a MN to impersonate a CN.

Also, it is important for the trusted entity CGA to use a high SEC value to avoid spoofing and thus compromise the entire RO operations. Futhermore, because its address must also be known globally, the trusted entity FQDN should either be hardcoded (for example ha.domain.com) or be a record inside the trusted zone. This prevents a MN to use any valid address from the same trusted domain to impersonate the trusted entity.

For an intruder that is not in a trusted domain nor colludes with any node in such domain, it must not only spoof a ECGA that is linked to a trusted domain, but will need the signed home and visited tokens to match its own generated ECGA parameters in order to execute a valid RO. These signatures can only be used if there is a hash collision with, respectively, the H(CGHoA | CGHoA Parameters) and H(CGCoA | CGCoA Parameters). The probability of such collision is $2^{-length\ of\ hash}$ and is therefore inversely proportional to the length of the hash.

A complete comparison of SMIPv6 with existing solutions is available in Table 6.6.

6.7 Conclusion

This paper proposes Secure MIPv6 (SMIPv6) that includes a new secure and efficient RO using an enhanced version of CGA (ECGA) and DNSSEC for MIPv6 to tackle the RR weak-

Table 6.6 Comparison of existing route optimisation protocols and the proposed SMIPv6 solution

Vulnerabilities and limitations	RR	CGA-OMIPv6	CAM-DH	RFC 4866	CBU	HCBU	SMIPv6
Session hijacking	X						
Man-in-the-middle	X	Δ	X	Δ	X		
Spoofing	X	Δ	Δ	Δ	Δ		
Replay	Δ	Δ	Δ	Δ			
Redirection	X	Δ	X	Δ	X		
Flooding	X	X	X	X	X	X	
Resource exhaustion DoS		Δ	X	Δ	Δ	Δ	Δ
Rely on nonexisting architectures						X	
High radio signalling overhead	Δ	X	Δ	X		Δ	
Nonexisting MN-CN reachability tests					X	X	
Scalability and flexibility issues			Δ		X		

Δ: Partially vulnerable

nesses and the unrealistic assumptions of existing works. First, ECGA integrates a backward key chain in the HASH-1 computation which makes it resilient to replay and time-memory tradeoff attacks while increasing the impersonation bruteforce complexity and still be very efficient in its generation and verification. Second, by combining ECGA and DNSSEC, SMIPv6 removes the burden of managing certificates, and thus improves the flexibility and scalability for network operators that only need to maintain a list of trusted domains. Moreover, the use of trusted domains enables the HA, VHA and CN to strongly authenticate all incoming messages and thus tackling all major attacks while relying on realistic assumptions and keeping the radio signalling overhead to a minimum. Results using AVISPA prove that the proposition is safe and does not contain any security flaw.

```
role session (
  MN, HA, DNS, VHA, CN : agent,  % Agents declaration
  PSK_MNHA : symmetric_key,      % PSK in MN's SIM card
  PSK_HADNS : symmetric_key,     % Access for managing home domain (zone)
  PSK_VHADNS : symmetric_key,    % Access for managing visited domain (zone)
  PK_HA : public_key,    % Public Key of HA
  PK_VHA : public_key,   % Public Key of visited HA
  PK_DNS : public_key,   % Public Key of DNS server
  HA_Domain : text,      % Zone/domain of the MN's home network
  VHA_Domain : text,     % Zone/domain of the MN's visited network
  HA_FQDN : text,        % FQDN of trusted entity in the MN's home network
  VHA_FQDN : text,       % FQDN of trusted entity in the MN's visited network
  PRF : hash_func) def= % Hash function used for ECGA and signatures


  local
      S,R : channel (dy)

  composition
       homeagent (MN, HA, DNS, VHA, CN, S, R, PRF, HA_Domain, HA_FQDN,
VHA_FQDN, PSK_MNHA, PSK_HADNS, PK_HA, PK_DNS)
      /\ mobilenode (MN, HA, DNS, VHA, CN, S, R, PRF, PSK_MNHA)
      /\ dnsserver (MN, HA, DNS, VHA, CN, S, R, PRF, HA_Domain, VHA_Domain,
HA_FQDN, VHA_FQDN, PSK_HADNS, PSK_VHADNS, PK_DNS)
      /\ visitedagent (MN, HA, DNS, VHA, CN, S, R, PRF, VHA_Domain, HA_FQDN,
VHA_FQDN, PSK_VHADNS, PK_VHA, PK_DNS)
      /\ correspondantnode(MN,HA,DNS,VHA,CN, S, R, PRF, PK_DNS, HA_Domain,
VHA_Domain, HA_FQDN, VHA_FQDN)
```

Figure 6.12 HLPSL specification of the role sessions

```
role environment() def=

   intruder_knowledge = kiha, kidns, i_domain, i_fqdn, pki,prf,pkha, pkdns, ha_domain,
vha_domain, ha_fqdn, vha_fqdn

  composition
        session(mn, ha, dns, vha, cn, kmnha, khadns, kvhadns, pkha, pkvha, pkdns,
ha_domain, vha_domain, ha_fqdn, vha_fqdn, prf)
       /\ session(i, ha, dns, vha, cn, kiha, khadns, kvhadns, pkha, pkvha, pkdns, ha_domain,
vha_domain, ha_fqdn, vha_fqdn, prf)
       /\ session(mn, i, dns, vha, cn, kiha, kidns, kvhadns, pki, pkvha, pkdns, i_domain,
vha_domain, ha_fqdn, vha_fqdn, prf)
       /\ session(mn, ha, dns, i, cn, kmnha, khadns, kvhadns, pkha, pki, pkdns, ha_domain,
vha_domain, ha_fqdn, vha_fqdn, prf)
       /\ session(i, i, dns, i, cn, kmnha, khadns, kvhadns, pkha, pki, pkdns, ha_domain,
vha_domain, ha_fqdn, vha_fqdn, prf)

end role
```

Figure 6.13 HLPSL specification of role environment

```
% OFMC
% Version of 2006/02/13
SUMMARY
  SAFE
DETAILS
  BOUNDED_NUMBER_OF_SESSIONS

PROTOCOL
  C:\progra\ ~1\SPAN\testsuite\results\RO.if

GOAL
  as_specified

BACKEND
  OFMC

STATISTICS
  parseTime: 0.00s
  searchTime: 1.22s
  visitedNodes: 55 nodes
  depth: 6 plies
```

```
CL-ATSE
SUMMARY
  SAFE
DETAILS
  BOUNDED_NUMBER_OF_SESSIONS
  TYPED_MODEL

PROTOCOL
  C:\progra\ ~1\SPAN\testsuite\results\RO.if

GOAL
  As Specified

BACKEND
  CL-AtSe

STATISTICS
  Analysed : 19580 states
  Reachable : 978 states
  Translation: 11.45 seconds
  Computation: 0.11 seconds
```

Figure 6.14 Results from AVISPA for SMIPv6

CHAPITRE 7

DISCUSSION GÉNÉRALE

Ce chapitre porte sur une analyse globale des travaux accomplis dans cette thèse. Elle débute par revenir sur les questions posées dans le chapitre d'introduction dans le but de voir si les objectifs ont été atteints en discutant brièvement des résultats et de leur portée. Les aspects méthodologiques menant vers la conception et l'évaluation de performance des solutions proposées viennent conclure le chapitre.

7.1 Synthèse des travaux et rencontre des objectifs

Cette thèse a donné lieu à trois publications scientifiques, un article soumis et cinq brevets industriels protégeant les idées des chapitres 5 et 6. Le travail derrière chacun de ces livrables ciblait des objectifs spécifiques et il est donc important de se pencher sur l'atteinte de celles-ci.

Le système de détection d'intrusion présenté au chapitre 3 modélise bien mathématiquement le risque d'une attaque complice dépendamment du positionnement des nœuds et de leur degré de confiance. Une classification plus exhaustive des nœuds et une restriction plus sévère entre les transitions réduit les abus de seuils interclasses sans toutefois les éliminer complètement. En effet, les résultats montrent bien l'amélioration apportée au débit, en comparaison avec le IDS standard watchdog/pathrater ou encore simplement le protocole de routage DSR sans IDS, notamment dans un contexte d'un nombre important d'attaquants complices. Le premier objectif est donc rencontré avec succès malgré que certaines vulnérabilités sont toujours présentes mais exigent une coordination accrue entre les nœuds malicieux pour être exploitées. Un article sur le sujet a été publié dans la revue *International Journal of Computer Science and Network Security*.

Toujours dans les MANETs, la deuxième question relate de la conception d'un cadre de travail théorique pour inciter les nœuds à coopérer afin de respecter des critères de QoS pendant une période de temps spécifiée tout en supposant des conditions réalistes. L'article du chapitre 4 présente un modèle économique de fixation de prix incitant les nœuds à coopérer pour retransmettre les messages selon des critères de QoS. Ainsi, un protocole de routage simple a été conçu pour les applications en temps réel avec support de QoS dans lequel la source compense les nœuds intermédiaires avec de la monnaie virtuelle qui leur donne un pouvoir d'utilisation du réseau pour soit transmettre leurs propres messages ou accéder à des services distants. La modélisation des comportements rationnels est faite sous la forme d'un

jeu en considérant une incertitude due à l'information imparfaite circulant dans le réseau. Ce cadre de travail ne forme qu'une base sur lequel d'autres facteurs peuvent venir s'y greffer pour prédire le comportement des nœuds dans d'autres types d'applications. Les résultats montrent bien la performance attendu du réseau en terme de fluctuation des prix et de bris de contrats selon l'incertitude des coûts des autres joueurs. L'objectif issu de la seconde question a donc été atteint. L'article traitant le sujet a été soumis à la revue *IEEE Transactions on Mobile Computing*.

Le second volet de la thèse traite de la multidiffusion à laquelle la troisième question cherche à concevoir un protocole sécuritaire pour des applications vidéo à grande échelle telles IPTV supportant un grand nombre d'abonnés qui utilisent à la fois des unités mobiles (PDA) et fixes (télévision). Le protocole de distribution de clés pour la multidiffusion présenté au chapitre 5 se base sur LKH, qui est reconnu pour sa stabilité au niveau de la complexité par un balancement aisé de l'arbre, dans lequel une chaîne de hachage bidirectionnelle est inclue à chaque nœud logique de son arbre hiérarchique. La solution permet la configuration dynamique d'une période de grâce permettant de régénérer les clés perdues dépendamment du taux de réception des messages de mise-à-jour des clés de l'utilisateur. Le faible taux de données indéchiffrables sous un environnement à grande mobilité susceptible aux pertes de messages témoigne de l'efficacité impressionnante de la solution comparativement au LKH de base sans clés régénératrices. Le troisième objectif a donc été rencontré avec brio comme le prouve la publication dans la revue *Journal of Network and Systems Management* de l'éditeur *Springer* ainsi que les deux applications de brevet approuvées par Ericsson(P28200US: *Self-healing encryption keys* & P296309US: *Self-healing key distribution for LKH*).

La dernière problématique entoure la sécurisation de l'optimisation de route dans MIPv6 dans le but d'offrir une authentification forte en se basant sur des technologies existantes et en optant pour une gestion simplifiée de l'information pré-partagée. La recherche effectuée sur le sujet a mené à une nouvelle version du CGA, nommé ECGA, qui inclut une chaîne de hachage inversée pour empêcher le pré-calcul du hash-2. L'imbrication de plusieurs adresses CGA en cascade est également possible pour construire une hiérarchie de dépendance d'adresses qui force l'annulation du CGA si l'une des adresses imbriquées devient invalide. L'algorithme d'ECGA a ensuite été intégrée au nouveau protocole d'optimisation de route qui utilise DNSSEC pour assurer une authentification forte basée sur le nom des domaines de confiance. La solution proposée au chapitre 6 répond complètement à la dernière question posée comme en témoigne sa publication dans la revue *Journal of Telecommunications Management* de l'éditeur *Henry Stewart publications* ainsi que l'acceptation des 3 applications de brevet par Ericsson (P32688US: *Enhanced cryptographically generated addresses for secure route optimization in mobile Internet protocol* & P33327US: *Secure route optimization*

in mobile Internet protocol using trusted domain name servers & P33328US: *Cryptographically generated addresses using backward key chain for secure route optimization in mobile Internet protocol*). L'implémentation du protocole a été effectuée sous le vérificateur de protocoles cryptographiques dans le modèle formel AVISPA pour s'assurer de l'absence de toute vulnérabilité.

7.2 Méthodologie

La thèse traite de problématiques dont certaines sont plus théoriques (chapitres 3 et 4) et d'autres plus appliquées (chapitres 5 et 6). De plus, dépendamment des objectifs attendus, à savoir s'ils touchent la performance générale du réseau ou de s'assurer de l'absence de vulnérabilité, certains éléments dans la méthodologie utilisée pour concevoir et évaluer la performance des solutions proposées vont varier. Cependant, chaque évaluation de performance nécessite un plan d'expérimentation qui détaille soigneusement les facteurs et les sessions à étudier.

La modélisation mathématique a permis le calcul du risque d'attaques complices dans les MANETs et conjointement aux outils de la théorie des jeux, notamment les concepts issus de la compétition de Bertrand, le cadre de travail modélisant un protocole incitatif pour supporter des critères de QoS a pu être conçu.

L'évaluation de la performance diffère d'une solution à l'autre dépendamment du contexte de la conception et des objectifs fixés. Par exemple, l'implémentation du IDS du chapitre 3 et du protocole de distribution de clés de groupe pour la multidiffusion détaillé au chapitre 5 ont été effectués en C/C++ dans le simulateur de réseaux QualNet (2009). Les indices de performance peuvent alors être très détaillées étant donné la disponibilité d'une panoplie de métriques. D'un autre côté, l'objectif principal du protocole d'optimisation de route pour MIPv6 est d'être sécuritaire et de n'être vulnérable à aucune attaque. Ainsi, son implémentation a été effectué sous *Automated Validation of Internet Security Protocols and Applications*(AVISPA) Armando *et al.* (2005), un vérificateur de protocoles cryptographiques dans le modèle formel, qui valide la sécurité des échanges de messages en regard des objectifs d'authentification, d'intégrité et de confidentialité des messages. Finalement, dans un contexte plus théorique, MATLAB (2009) a été utilisé pour résoudre les systèmes d'équations différentielles partielles et trouver l'équilibre de Nash pour voir le comportement à long terme du réseau.

CHAPITRE 8

CONCLUSION ET RECOMMANDATIONS

Pour conclure cette thèse, les différentes contributions sont mises en évidence en mettant l'emphase sur les innovations apportées dans le domaine de la sécurité des réseaux informatiques mobiles. Les limitations suivront ensuite pour terminer avec les recommandations sur des voies de recherche ultérieures qui pourraient susciter de l'intérêt.

8.1 Contributions de la thèse

Cette thèse touche à la sécurité des réseaux mobiles ad hoc et cellulaires dans différents contextes d'application. Cette diversité a mené vers des contributions touchant un plus large éventail de problématiques mais toujours sous le thème générale d'améliorer la sécurité des réseaux mobiles de nouvelle génération. Plus précisément, la liste qui suit résume les principales contributions de la thèse.

1. Modélisation mathématique du risque d'attaques complices et son intégration dans le composant pathrater du système de détection d'intrusion pour les MANETs. Ainsi, le critère de sélection d'une route ne se limite pas simplement à la moyenne de la réputation des nœuds intermédiaires composant le chemin, mais intègre également le nombre de nœuds observateurs passifs pouvant surveiller les échanges de messages ainsi que leur réputation. Une classification plus exhaustive des nœuds dans le composant watchdog du IDS évite l'isolement des nœuds faussement accusés comme défaillant et réduit les risques d'abus des seuils interclasses sans toutefois l'éliminer complètement. Une comparaison avec DSR et la solution originale watchdog/pathrater a été effectuée.
2. L'utilisation des outils de la théorie des jeux a donné lieu à la modélisation du comportement rationnel des nœuds dans un protocole de routage simple permettant à la source d'émettre un contrat dans lequel ses critères de QoS à respecter pendant une certaine durée sont spécifiés. Le modèle se fonde sur les principes d'une compétition de Bertrand et considère des suppositions réalistes notamment au niveau de la qualité de l'information reçue par les nœuds et à la volonté d'un joueur de participer au marché selon le coût estimé des compétiteurs. La performance théorique du protocole en termes de taux de bris de contrat et de la tendance des prix selon l'incertitude des coûts des compétiteurs a également été montrée.

3. Conception d'un protocole de distribution de clés régénératrices pour LKH supportant des applications vidéos en temps-réel pour un très grand nombre d'abonnés sur des unités autant fixes que mobiles. La solution offre une performance impressionnante comparativement à LKH dans un environnement à haut taux de perte de paquets comme le montre l'évaluation de performance effectuée. Une analyse de la sécurité et de la complexité montrent l'absence de vulnérabilité et la stabilité des coûts de la mise-à-jour des clés avec un nombre grandissant d'abonnés.
4. Conception d'un algorithme de génération cryptographique d'adresse basé sur CGA qui inclut une chaîne de hachage inversée et supporte l'imbrication d'adresses en cascade pour fins d'éliminer l'attaque du compromis temps-mémoire et de monter une hiérarchie d'adresses jusqu'à l'atteinte d'une entité de confiance. Cette version améliorée de CGA est à la base de la solution du point suivant.
5. Conception d'un protocole d'optimisation de route pour MIPv6 qui, avec une simple gestion de domaines de confiance, assure une authentification forte avec l'utilisation de DNSSEC. Cette solution utilise une authentification décentralisée et utilise des technologies existantes aujourd'hui ou sur le point d'être déployées, contrairement aux solutions dans la littérature avec un même niveau élevé de sécurité.

8.2 Limitations

L'identification des limitations des travaux présentés est une étape importante à l'avancement de la science. Ainsi, pour chacun des quatre articles, les limites sont détaillées. Dans le chapitre 3, le calcul du chemin le plus fiable tient compte du nombre d'observateurs ainsi que de leur réputation qui sont en mesure de capter le message à transférer et celui à retransférer par le nœud intermédiaire. Cependant, la réputation d'un nœud observateur dépend de son comportement dans la retransmission des messages. Il est ainsi possible qu'il agisse d'une manière exemplaire dans sa tâche de relayeur de messages, mais qu'il mente sur les observations détectées. Un groupe de nœuds malicieux peut ainsi favoriser la sélection d'un chemin qui inclut des nœuds intermédiaires pouvant modifier les messages sans que les observateurs complices ne dénoncent l'action. De plus, l'abus des seuils intra-classe demeure toujours possible malgré la classification plus exhaustive des nœuds. Cette problématique est inhérente à tout système basé sur le crédit ou la réputation. Une autre limite, plus générale, est la surcharge de travail imposée aux nœuds intermédiaires qui sont fiables et entourés de nœuds observateurs légitimes. Les nœuds égoïstes sont ainsi avantagés par leur utilisation moindre de ressources énergétiques et en bande passante.

Tous les mécanismes incitatifs basés sur un modèle économique nécessitent un matériel

infraudable pouvant assurer une gestion intègre de la monnaie virtuelle. Dans un contexte décentralisé comme la solution proposée au chapitre 4, l'utilisation de la carte à puce devient essentielle. Quoique cette technologie existe depuis fort longtemps, son adoption dans les unités mobiles n'est pas assurée. Il aurait de plus été intéressant d'inclure dans le modèle d'autres métriques importantes de QoS pour les applications en temps réel telles le délai de bout en bout ou encore le jitter. Finalement, la conversion du modèle en un jeu dynamique ou répété aurait ajouté davantage de réalisme étant donnée l'apprentissage faits par les nœuds selon les observations effectuées au cours du temps.

Comme décrit au chapitre 5, l'utilisation d'une chaîne de hachage bidirectionnelle dans le protocole de distribution de clés régénératrices présente des risques de collusion entre des usagers (ou un même usager) qui ont été abonnés à différents moments afin de déchiffrer le trafic entre le premier départ et la seconde arrivée au groupe. La deuxième version du protocole (*scheme 2*) règle cette problématique, mais vient au détriment de la performance lors de la révocation d'usagers. Le meilleur compromis est donc d'opter pour une révocation groupée (*batched revocation*).

La principale limite du protocole d'optimisation de route présenté au chapitre 6 est de forcer le réseau visité à suivre le même protocole que le réseau mère. Or dans un contexte où de nombreux opérateurs existent au niveau mondial, imposer une telle condition requiert un consortium commun. Si un protocole visité ne supporte pas l'optimisation de route proposée, la communication tunnelisée devient la seule option sécuritaire. De plus, il aurait été intéressant d'implémenter la solution dans un simulateur dans le but d'évaluer sa performance notamment au niveau des délais.

8.3 Suggestions de travaux futurs

Suite aux limitations identifiées, d'autres avenues de recherche parallèles peuvent être intéressantes à entreprendre. Cette thèse conclut avec la présentation de quelques recommandations.

L'objectif du protocole de distribution de clé proposé pour la multidiffusion est de supporter un très grand nombre d'abonnés tout en ayant une complexité relativement stable avec le nombre d'arrivées et de départs, et en offrant un bon taux de déchiffrement du trafic malgré les pertes de messages de mise-à-jour des clés. Or, une alternative intéressante serait d'utiliser SDR comme base et de proposer des extensions permettant à ce protocole d'offrir une complexité stable indépendamment des actions des usagers. L'avantage de SDR sur LKH est qu'une extension de clés régénératrices existe déjà et n'est pas sensible au départ des abonnées.

En second lieu, il serait intéressant considérer la sécurisation de l'optimisation de route dans Proxy Mobile IPv6 (PMIPv6) qui est une version de MIPv6 où les tâches de l'unité mobile sont léguées au réseau des opérateurs. Quelques travaux traitent déjà du sujet B. Sarikaya (2008); Song *et al.* (2009); Han *et al.* (2008), mais aucune solution ressort du lot en offrant une authentification forte et une gestion simplifiée tout en étant efficace.

Pour terminer, le modèle économique présenté au chapitre 4 offre une bonne base théorique afin de l'étendre pour y intégrer de nouvelles métriques de QoS selon les applications désirées pour les MANETs. Il serait donc intéressant d'exploiter ce potentiel en considérant, par exemple, des critères de QoS plus souples et donc tolérants au bris ou défaillance temporaire. Également, une étude pourrait être portée sur la période de temps dynamique à laquelle les nœuds intermédiaires sont payés. Ainsi, au lieu de limiter les paiements uniquement à l'échéance du contrat, il aurait peut-être été plus profitable pour la source de considérer une durée de contrat plus petite dans le cas où par exemple plusieurs routes alternatives sont disponibles. Aussi, pour les applications dont l'ordre des messages n'est pas une priorité, le modèle pourrait être étendu pour diviser un même contrat à travers plusieurs joueurs. Finalement, comme mentionnée dans la section 8.2, l'ajout du dynamisme au modèle, où les actions précédentes d'un joueur ont un impact sur les décisions à venir, serait certainement un atout important malgré la complexité mathématique qu'il pourrait engendrer.

RÉFÉRENCES

(2011). Dns whitelist: Protect against false positives. `http://www.dnswl.org/`. [Online; accessed 20-February-2011].

AFERGAN, M. (2006). Using repeated games to design incentive-based routing systems. *INFOCOM 2006. 25th IEEE International Conference on Computer Communications. Proceedings*, 1–13.

ANDEREGG, L. et EIDENBENZ, S. (2003). Ad hoc-vcg: a truthful and cost-efficient routing protocol for mobile ad hoc networks with selfish agents. *MobiCom '03: Proceedings of the 9th annual international conference on Mobile computing and networking*. ACM, New York, NY, USA, 245–259.

ARENDS, R., AUSTEIN, R., LARSON, M., MASSEY, D. et ROSE, S. (2005a). DNS Security Introduction and Requirements. RFC 4033 (Proposed Standard).

ARENDS, R., AUSTEIN, R., LARSON, M., MASSEY, D. et ROSE, S. (2005b). Protocol Modifications for the DNS Security Extensions. RFC 4035 (Proposed Standard). Updated by RFC 4470.

ARENDS, R., AUSTEIN, R., LARSON, M., MASSEY, D. et ROSE, S. (2005c). Resource Records for the DNS Security Extensions. RFC 4034 (Proposed Standard). Updated by RFC 4470.

ARKKO, J., VOGT, C. et HADDAD, W. (2007). Enhanced Route Optimization for Mobile IPv6. RFC 4866 (Proposed Standard).

ARMANDO, A., BASIN, D., BOICHUT, Y., CHEVALIER, Y., COMPAGNA, L., CUELLAR, J., DRIELSMA, P. H., HEÁM, P. C., KOUCHNARENKO, O., MANTOVANI, J., MÖDERSHEIM, S., VON OHEIMB, D., RUSINOWITCH, M., SANTIAGO, J., TURUANI, M., VIGANÒ, L. et VIGNERON, L. (2005). The avispa tool for the automated validation of internet security protocols and applications. *Computer Aided Verification*,

Springer Berlin Heidelberg, Berlin, Heidelberg, vol. 3576, chapitre 27. 281–285.

AURA, T. (2003). Cryptographically generated addresses (cga). *Information Security*, Springer Berlin / Heidelberg, vol. 2851 de *Lecture Notes in Computer Science*. 29–43.

B. SARIKAYA, A. QIN, A. H. W. W. (2008). *RFCDRAFT: PMIPv6 Route Optimization Protocol*. IETF. Status: DRAFT revision 0.

BAYE, M. R. et KOVENOCK, D. (2008). Bertrand competition. S. N. Durlauf et L. E. Blume, éditeurs, *The New Palgrave Dictionary of Economics*, Palgrave Macmillan, Basingstoke.

BLUME, A. (2003). Bertrand without fudge. *Economics Letters*, 78, 167–168.

BLUNDO, C., D'ARCO, P., DE SANTIS, A. et STINSON, D. (2007). On unconditionally secure distributed oblivious transfer. *Journal of Cryptology*, 20, pp. 323 – 73.

BOS, J., OZEN, O. et HUBAUX, J.-P. (2009). Analysis and optimization of cryptographically generated addresses. P. Samarati, M. Yung, F. Martinelli et C. Ardagna, éditeurs, *Information Security*, Springer Berlin / Heidelberg, vol. 5735 de *Lecture Notes in Computer Science*. 17–32.

BUCHEGGER, S. et BOUDEC, J.-Y. L. (2002). Performance analysis of the confidant protocol. *MobiHoc '02: Proceedings of the 3rd ACM international symposium on Mobile ad hoc networking & computing*. ACM, New York, NY, USA, 226–236.

CAPKUN, S., BUTTYAN, L. et HUBAUX, J.-P. (2003). Self-organized public-key management for mobile ad hoc networks. *Mobile Computing, IEEE Transactions on*, 2, 52–64.

CHEN, W. et DONDETI, L. (2003a). Recommendations in using group key management algorithms. *DARPA Information Survivability Conference and Exposition, 2003. Proceedings*. vol. 2, 222–227 vol.2.

CHEN, W. et DONDETI, L. (2003b). Recommendations in using group key management algorithms. *DARPA Information Survivability Conference and Exposition, 2003. Proceedings*. vol. 2, pp. 222–227.

CHO, T., LEE, S.-H. et KIM, W. (2004). A group key recovery mechanism based on logical key hierarchy. *Journal of Computer Security*, 12, pp. 711 – 36.

COOK, J. (January 2009). Mobility and reliability modeling for a mobile ad hoc network. *IEEE Transactions*, 41, 23–31(9).

D., W., E., H. et R., A. (1999). Key management for multicast: Issues and architectures. RFC 2627, Internet Engineering Task Force.

D. JOHNSON, D. MALTZ, Y.-C. H. (2007). *RFCDRAFT: The Dynamic Source Routing Protocol for Mobile Ad Hoc Networks (DSR)*. IETF. Status: DRAFT revision 10.

D. KAVITHA, K.E.SREENIVASA MURTHY, B. S. V. R. S. U. H. (2010). An efficient hierarchical certificate based binding update protocol for route optimization in mobile ipv6. *Global Journal of Computer Science and Technology*, 10.

DASILVA, L. A., PETR, D. W. et AKAR, N. (2000). Static Pricing and Quality of Service in Multiple Service Networks. *Proc. 5th Joint Conf. Information Sciences, Atlantic City, NJ*. vol. 1, 355–358.

DASILVA, L. A. et SRIVASTAVA, V. (2004). Node participation in ad-hoc and peer-to-peer networks: A game-theoretic formulation. *In Proc. Workshop on Games and Emergent Behavior in Distributed Computing Environments*, 11, 21 – 38.

DASTIDAR, K. G. (1995). On the existence of pure strategy bertrand equilibrium. *Economic Theory*, 5, 19–32.

DENG, R. H., ZHOU, J. et BAO, F. (2002). Defending against redirect attacks in mobile ip. *Proceedings of the 9th ACM conference on Computer and communications security*. ACM, New York, NY, USA, CCS '02, 59–67.

DRIELSMA, P. H., MÖDERSHEIM, S., VIGANÒ, L. et BASIN, D. (2007). Formalizing and analyzing sender invariance. *FAST'06: Proceedings of the 4th international conference on Formal aspects in security and trust*. Springer-Verlag, Berlin, Heidelberg, 80–95.

DUTTA, R., CHANG, E.-C. et MUKHOPADHYAY, S. (2007). Efficient self-healing key distribution with revocation for wireless sensor networks using one way key chains. pp. 385 – 400.

DUTTA, R., WU, Y. D. et MUKHOPADHYAY, S. (2008). Constant storage self-healing key distribution with revocation in wireless sensor network. pp. 1323 – 8.

EIDENBENZ, S., RESTA, G. et SANTI, P. (2008). The commit protocol for truthful and cost-efficient routing in ad hoc networks with selfish nodes. *Mobile Computing, IEEE Transactions on*, 7, 19–33.

GHOSH, T., PISSINOU, N. et MAKKI, K. (2004). Collaborative trust-based secure routing against colluding malicious nodes in multi-hop ad hoc networks. *Local Computer Networks, 2004. 29th Annual IEEE International Conference on*. 224 – 231.

GHOSH, T., PISSINOU, N. et MAKKI, K. (2005). Towards designing a trusted routing solution in mobile ad hoc networks. *Mob. Netw. Appl.*, 10, 985–995.

GUNDAVELLI, S., LEUNG, K., DEVARAPALLI, V., CHOWDHURY, K. et PATIL, B. (2008). Proxy Mobile IPv6. RFC 5213 (Proposed Standard).

HAN, B.-J., LEE, J.-M., LEE, J.-H. et CHUNG, T.-M. (2008). Pmipv6 route optimization mechanism using the routing table of mag. *Proceedings of the 2008 Third International Conference on Systems and Networks Communications*. IEEE Computer Society, Washington, DC, USA, 274–279.

HOERNIG, S. H. (2002). Mixed bertrand equilibria under decreasing returns to scale: an embarrassment of riches. *Economics Letters*, 74, 359–362.

HONG, D. et KANG, J.-S. (2005). An efficient key distribution scheme with self-healing property. *Communications Letters, IEEE*, 9, pp. 759–761.

HU, Y.-C., PERRIG, A. et JOHNSON, D. (2005). Ariadne: a secure on-demand routing protocol for ad hoc networks. *Wireless Networks*, 11, 21 – 38.

IOANNIDIS, J., KEROMYTIS, A. D. et YUNG, M., éditeurs (2005). *Applied Cryptography and Network Security, Third International Conference, ACNS 2005, New York, NY, USA, June 7-10, 2005, Proceedings*, vol. 3531 de *Lecture Notes in Computer Science*.

JANSSEN, M. et RASMUSEN, E. (2002). Bertrand competition under uncertainty. *Journal of Industrial Economics*, 50, 11–21.

JARAMILLO, J. J. et SRIKANT, R. (2007). Darwin: distributed and adaptive reputation mechanism for wireless ad-hoc networks. *MobiCom '07: Proceedings of the 13th annual ACM international conference on Mobile computing and networking*. ACM, New York, NY, USA, 87–98.

JOHNSON, D., PERKINS, C. et ARKKO, J. (2004). Mobility Support in IPv6. RFC 3775 (Proposed Standard).

KAPLAN, T. R. et WETTSTEIN, D. (2000). The possibility of mixed-strategy equilibria with constant-returns-to-scale technology under bertrand competition. *Spanish Economic Review*, 2, 65–71.

KAUSAR, F., HUSSAIN, S., PARK, J. H. et MASOOD, A. (2007). Secure group communication with self-healing and rekeying in wireless sensor networks. pp. 737 – 48.

KAVITHA, D., MURTHY, K. et UL HUQ, S. (2010). Security analysis of binding update protocols in route optimization of mipv6. Piscataway, NJ, USA, 44 – 9.

KEMPF, J. et KOODLI, R. (2008). Distributing a Symmetric Fast Mobile IPv6 (FMIPv6) Handover Key Using SEcure Neighbor Discovery (SEND). RFC 5269 (Proposed Standard).

LAZAR, A., ORDA, A. et PENDARAKIS, D. (1997). Virtual path bandwidth allocation in multiuser networks. *Networking, IEEE/ACM Transactions on*, 5, 861–871.

LIU, D., NING, P. et SUN, K. (2003). Efficient self-healing group key distribution with revocation capability. *Proc. of the 10th ACM Conference on Computer and Communications Security (CCS '03)*.

LIU, Z., JOY, A. W. et THOMPSON, R. A. (2004). A dynamic trust model for mobile ad hoc networks. *Future Trends of Distributed Computing Systems, IEEE International Workshop*, 0, 80–85.

LIZHI WANG, MAINAK MAZUMDAR, M. D. B. et VALENZUELA, J. (2007). Oligopoly models for market price of electricity under demand uncertainty and unit reliability. *European Journal of Operational Research*, 181, 1309–1321.

LU, B. et POOCH, U. (2004). A game theoretic framework for bandwidth reservation in mobile ad hoc networks. *Quality of Service in Heterogeneous Wired/Wireless Networks, 2004. QSHINE 2004. First International Conference on*, 234–241.

M. ROE, T. AURA, G. O. J. A. (2002). Authentication of Mobile IPv6 Binding Updates and Acknowledgments. IETF Internet Draft.

MACKENZIE, A. B. et DASILVA, L. A. (2006). *Game Theory for Wireless Engineers (Synthesis Lectures on Communications)*. Morgan & Claypool Publishers.

MAHAJAN, V., NATU, M. et SETHI, A. (2008). Analysis of wormhole intrusion attacks in manets. *Military Communications Conference, 2008. MILCOM 2008. IEEE*. 1 –7.

MARQUEZ, R. (1997). A note on bertrand competition with asymmetric fixed costs. *Economics Letters*, 57, 87–96.

MARSHALL, J., THAKUR, V. et YASINSAC, A. (2003). Identifying flaws in the secure routing protocol. *Performance, Computing, and Communications Conference, 2003. Conference Proceedings of the 2003 IEEE International*. 167 – 174.

MARTI, S., GIULI, T. J., LAI, K. et BAKER, M. (2000). Mitigating routing misbehavior in mobile ad hoc networks. *MobiCom '00: Proceedings of the 6th annual international conference on Mobile computing and networking*. ACM, New York, NY, USA, 255–265.

MATLAB (2009). *version 7.9.0 (R2009b)*. The MathWorks Inc., Natick, Massachusetts.

MICHIARDI, P. et MOLVA, R. (2002). Core: a collaborative reputation mechanism to enforce node cooperation in mobile ad hoc networks. *Proceedings of the IFIP TC6/TC11 Sixth Joint Working Conference on Communications and Multimedia Security*. Kluwer, B.V., Deventer, The Netherlands, The Netherlands, 107–121.

MICHIARDI, P. et MOLVA, R. (2003). A game theoretical approach to evaluate cooperation enforcement mechanisms in mobile ad hoc networks. *In Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks*. 3–5.

MIYAO, K., NAKAYAMA, H., ANSARI, N., NEMOTO, Y. et KATO, N. (2008). A reliable topology for efficient key distribution in ad-hoc networks. *World of Wireless, Mobile and Multimedia Networks, 2008. WoWMoM 2008. 2008 International Symposium on a*, 1–5.

MORE, S. M., MALKIN, M., STADDON, J. et BALFANZ, D. (2003). Sliding-window self-healing key distribution. *SSRS '03: Proceedings of the 2003 ACM workshop on Survivable and self-regenerative systems*. pp. 82–90.

NAOR, D., NAOR, M. et LOTSPIECH, J. B. (2001). Revocation and tracing schemes for stateless receivers. *Proceedings of the 21st Annual International Cryptology Conference on Advances in Cryptology*. Springer-Verlag, London, UK, CRYPTO '01, 41–62.

NG, K. S. et SEAH, W. (2003). Routing security and data confidentiality for mobile ad hoc networks. *Vehicular Technology Conference, 2003. VTC 2003-Spring. The 57th IEEE Semiannual*, 3, 1821–1825 vol.3.

NIKANDER, P., ARKKO, J., AURA, T., MONTENEGRO, G. et NORDMARK, E. (2005). Mobile IP Version 6 Route Optimization Security Design Background. RFC 4225 (Informational).

PADRO, C., SÁEZ, G. et VILLAR, J. L. (1999). Detection of cheaters in vector space secret sharing schemes. *Des. Codes Cryptography*, 16, pp. 75–85.

PERKINS, C. (2002). IP Mobility Support for IPv4. RFC 3344 (Informational).

PERRIG, A., SONG, D. et TYGAR, J. (2001). Elk, a new protocol for efficient large-group key distribution. pp. 247 – 262.

QIU, Y. et MARBACH, P. (2003). Bandwidth allocation in ad hoc networks: a price-based approach. *INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications Societies. IEEE*, 2, 797–807 vol.2.

QUALNET (2009). *version 4.0*. Scalable Networks Technologies, Inc., Los Angeles, California.

RAJ, S. B. E. et LALITH, J. J. (2009). A novel approach for computation-efficient rekeying for multicast key distribution. *IJCSNS International Journal of Computer Science and Network Security*. vol. 9, pp. 279–284.

REBAHI, Y., MUJICA, V. et SISALEM, D. (2005). A reputation-based trust mechanism for ad hoc networks. *Proceedings of the 10th IEEE Symposium on Computers and Communications*. IEEE Computer Society, Washington, DC, USA, 37–42.

REN, K., LOU, W., ZENG, K., BAO, F., ZHOU, J. et DENG, R. H. (2006). Routing optimization security in mobile ipv6. *Computer Networks*, 50, 2401 – 2419.

ROUTLEDGE, R. R. (2010). Bertrand competition with cost uncertainty. *Economics Letters*, 107, 356–359.

S. BRADNER, A. MANKIN, J. S. (2003). A Framework for Purpose-Built Keys (PBK). IETF Internet Draft.

SETIA, S., ZHU, S. et JAJODIA, S. (2002). A comparative performance analysis of reliable group rekey transport protocols for secure multicast. vol. 49, pp. 21 – 41.

SHERMAN, A. T. et MCGREW, D. A. (2003). Key establishment in large dynamic groups using one-way function trees. *IEEE Trans. Softw. Eng.*, 29, pp. 444–458.

SHI, M., SHEN, X., JIANG, Y. et LIN, C. (2007). Self-healing group-wise key distribution schemes with time-limited node revocation for wireless sensor networks. *IEEE Wireless Communications*, 14, pp. 38 – 46.

SOLIMAN, H., CASTELLUCCIA, C., MALKI, K. E. et BELLIER, L. (2005). Hierarchical Mobile IPv6 Mobility Management (HMIPv6). RFC 4140 (Experimental). Obsoleted by RFC 5380.

SONG, J., KIM, H. et HAN, S. (2009). Route optimization in pmipv6 environment. *Computer and Information Technology, 2009. CIT '09. Ninth IEEE International Conference on*. vol. 2, 341 –346.

SPULBER, D. F. (1995). Bertrand competition when rivals' costs are unknown. *Journal of Industrial Economics*, 43, 1–11.

SRIVASTAVA, V., NEEL, J., MACKENZIE, A., MENON, R., DASILVA, L., HICKS, J., REED, J. et GILLES, R. (2005). Using game theory to analyze wireless ad hoc networks. *Communications Surveys & Tutorials, IEEE*, 7, 46–56.

STADDON, J., MINER, S., FRANKLIN, M., BALFANZ, D., MALKIN, M. et DEAN, D. (2002). Self-healing key distribution with revocation. pp. 241 – 257.

STERNE, D., LAWLER, G., GOPAUL, R., RIVERA, B., MARCUS, K. et KRUUS, P. (2007). Countering false accusations and collusion in the detection of in-band wormholes. *Computer Security Applications Conference, 2007. ACSAC 2007. Twenty-Third Annual*. 243 –256.

SU, X. et BOPPANA, R. (2007). On mitigating in-band wormhole attacks in mobile ad hoc networks. *Communications, 2007. ICC '07. IEEE International Conference on*. 1136 –1141.

SURESH P., L., KAUR, R., GAUR, M. S. et LAXMI, V. (2010). A collusion attack detection method for olsr-based manets employing scruple packets. *Proceedings of the 3rd international conference on Security of information and networks*. ACM, New York, NY, USA, SIN '10, 256–262.

TIAN, B., HAN, S., DILLON, T. S. et DAS, S. (2008). A self-healing key distribution scheme based on vector space secret sharing and one way hash chains.

TISSIER, P. E. (1984). Bertrand's Paradox. *The Mathematical Gazette*, 68, 15–19.

VIVES, X. (1999). *Oligopoly Pricing: Old Ideas and New Tools*. MIT Press.

W. HADDAD, L. MADOUR, J. A. F. D. (2005). Applying Cryptographically Generated Addresses to Optimize MIPv6 (CGA-OMIPv6). IETF Internet Draft.

W. HADDAD, M. N. (2009). On Using 'Symbiotic Relationship' to Repel Network Flooding Attack. IETF Internet Draft.

WEIFENG CHEN, LAKSHMINATH R. DONDETI, Y. S. (2008). Performance comparison of stateful and stateless group rekeying algorithms. *IJCSNS International Journal of Computer Science and Network Security*. vol. 8.

WONG, C. K., GOUDA, M. et LAM, S. S. (2000). Secure group communications using key graphs. *IEEE/ACM Trans. Netw.*, 8, pp. 16–30.

XIAO, Y., SHAN, X. et REN, Y. (2005). Game theory models for ieee 802.11 dcf in wireless ad hoc networks. *Communications Magazine, IEEE*, 43, S22–S26.

XUE, Y., LI, B. et NAHRSTEDT, K. (2003). Price-based resource allocation in wireless ad hoc networks. *in Proceedings of the Eleventh International Workshop on Quality of Service (IWQoS 2003), also Lecture Notes in Computer Science, ACM*. Springer-Verlag, 79–96.

YOO, Y. et AGRAWAL, D. P. (2006). Why does it pay to be selfish in a manet? *Wireless Communications, IEEE*, 13, 87–97.

ZHANG, X., LAM, S., LEE, D.-Y. et YANG, Y. (2003). Protocol design for scalable and reliable group rekeying. *IEEE/ACM Transactions on Networking*, 11, pp. 908 – 22.

ZHONG, S., CHEN, J. et YANG, Y. (2003). Sprite: a simple, cheat-proof, credit-based system for mobile ad-hoc networks. *INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications Societies. IEEE*, 3, 1987–1997 vol.3.

ZHOU, C., QIAN, D. et LEE, H. (2004). Utility-based routing in wireless ad hoc networks. *Mobile Ad-hoc and Sensor Systems, 2004 IEEE International Conference on*, 588–593.

ZHU, S. et JAJODIA, S. (2003). Scalable group rekeying for secure multicast: a survey. pp. 1 – 10.

ZHU, S., SETIA, S. et JAJODIA, S. (2003). Adding reliable and self-healing key distribution to the subset difference group rekeying method. *In Group Communications and Charges: Technology and Business Models. Proceedings of the 5th COST 264 International Workshop on Networked Group Communications, NGC 2003*. pp. 107–118.